研究生课程



数字图像处理
Digital Image Processing

彭宇新

北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

数字图像处理



- 课程介绍
  - ✓ 教材及参考书
  - ✓ 上课主要内容
  - ✓ 课程目的与要求
  - ✓ 考试内容
  - ✓ 助教老师
  - ✓ 国内外相关会议和杂志
  - ✓ 目前需要做的事情

教材及参考书



- 教材

  ✓ Rafael C. Gonzalez, Richard E. Woods 著，阮秋琦、阮宇智等译，数字图像处理（第二版），电子工业出版社，2003年。

  ✓ Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing(Second Edition), Prentice Hall, 2003。

教材及参考书（续）



- 参考书
  - ✓ Kenneth R. Castleman著，朱志刚、林学訚、石定机等译，数字图像处理，电子工业出版社，2002年。
  - ✓ 章毓晋，图象工程上册—图象处理和分析，清华大学出版社，2003年。
  - ✓ 阮秋琦，数字图像处理学，电子工业出版社，2004年。
  - ✓ 杨枝灵、王开等，Visual C++数字图像获取、处理及实践应用，人民邮电出版社，2003年。
  - ✓ 章毓晋，基于内容的视觉信息检索，科学出版社，2003年。

上课主要内容



1. 概述
2. 空间域图像增强
3. 彩色图像处理
4. 基于内容的图像检索（补充）
5. 傅里叶变换
6. 频率域图像增强
7. 图像复原
8. 图像压缩（1）

上课主要内容（续）



9. 图像压缩（2）
10. 形态学图像处理
11. 图像分割
12. 表示与描述
13. 基于内容的视频分析和检索技术（补充）
14. 考试复习

课程目的与要求



- 掌握数字图像处理的基本概念、原理和方法
- 初步运用所学知识解决实际问题
- 为图像处理及相关领域的研究打下基础
  - ✓ 图像处理
  - ✓ 计算机视觉
  - ✓ 基于内容的图像、视频检索
  - ✓ 人脸识别、指纹识别、掌纹识别、虹膜识别
  - ✓ 图像分类、图像和视频的语义概念检测、……

考试内容



- 平时作业50%，考试成绩50%
- 平时作业选择下列两个之一：
  - 分组完成一个大作业
  - 其他与图像处理相关的课题(需要上课老师认可)
- 考试闭卷完成（基本概念、原理和算法）

助教老师



- 曹磊

  E_mail: caolei@icst.pku.edu.cn

  电　话：82529384

注意：有不懂的地方多问助教老师，充分发挥助教老师的指导作用



国内外相关会议和杂志



- 国内杂志：一级学报
  - ✓ Journal of Computer Science and Technology (JCST)(SCI)
  - ✓ 计算机学报
  - ✓ 软件学报
  - ✓ 电子学报
  - ✓ 计算机研究与发展
  - ✓ …

国内外相关会议和杂志（续）



- 国外会议：
  - IEEE International Conference on Computer Vision (ICCV)
  - IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)
  - ACM Multimedia Conference (MM)





- 国外会议：
  - ✓ IEEE International Conference on Image Processing (ICIP)
  - ✓ IEEE International Conference on Multimedia and Expo (ICME)
  - ✓ International Conference on Pattern Recognition (ICPR)
  - ✓ ACM International Conference on Image and Video Retrieval (CIVR)
  - ✓ …



国内外相关会议和杂志（续）



- 国外期刊：
  - ✓ IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)
  - ✓ IEEE Transactions on Image Processing (IP)
  - ✓ IEEE Transactions on Circuits and Systems for Video Technology (CSVT)
  - ✓ International Journal of Computer Vision (IJCV)
  - ✓ Pattern Recognition (PR)
  - ✓ Image and Vision Computing (IVC)
  - ✓ …

目前需要做的事情



- 选课学生发送下列信息给老师：pengyuxin@icst.pku.edu.cn
  - 姓名
  - 学号
  - 联系方式：E_mail，电话
  - 硕士生或博士生，年级
  - 所在院系、实验室、导师
  - 研究方向

数字图像处理基础



- 概述
  - ✓ 概念：图像、数字图像、像素
  - ✓ 数字图像处理的起源
  - ✓ 数字图像处理的应用领域
  - ✓ 图像处理系统的部件
- 基础知识
  - ✓ 图像的采样和量化
  - ✓ 数字图像的表示
  - ✓ 数字图像的质量
  - ✓ 像素间的一些基本关系

什么是图像?



✓定义为二维函数f(x, y), 其中， x, y是空间坐标， f(x, y)是点（x, y）的幅值

✓灰度图像是一个二维灰度（或亮度）函数f(x, y)

✓彩色图像由三个（如RGB, HSV）二维灰度（或亮度）函数f(x, y)组成





什么是数字图像？



- ✓ 像素组成的二维排列，可以用矩阵表示

- ✓ 对于单色（灰度）图像而言，每个像素的亮度用一个数值来表示，通常数值范围在0到255之间，0表示黑、255表示白，其它值表示处于黑白之间的灰度

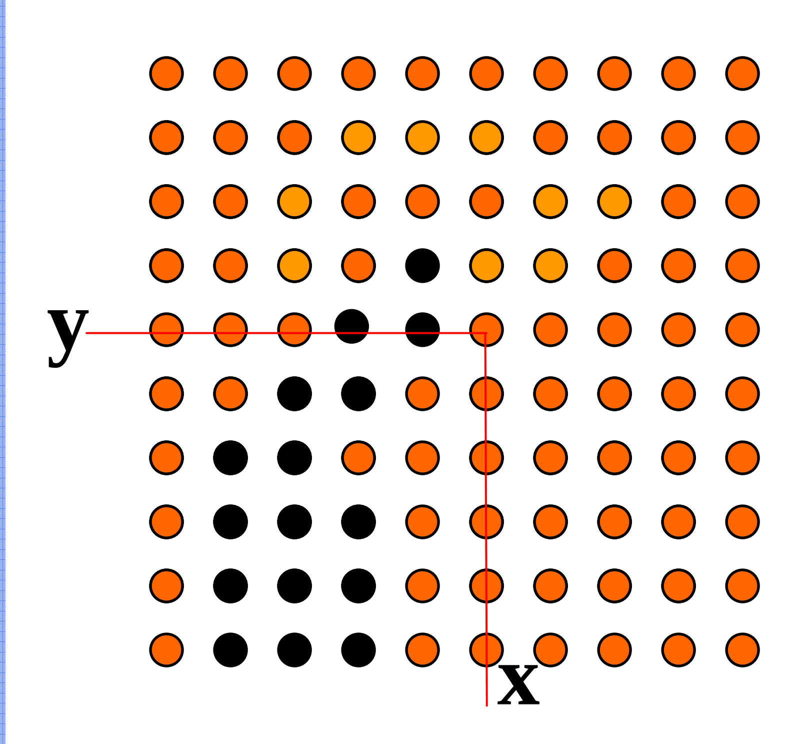
- ✓ 彩色图像可以用红、绿、蓝三元组的二维矩阵来表示。通常，三元组的每个数值也是在0到255之间，0表示相应的基色在该像素中没有，而255则代表相应的基色在该像素中取得最大值

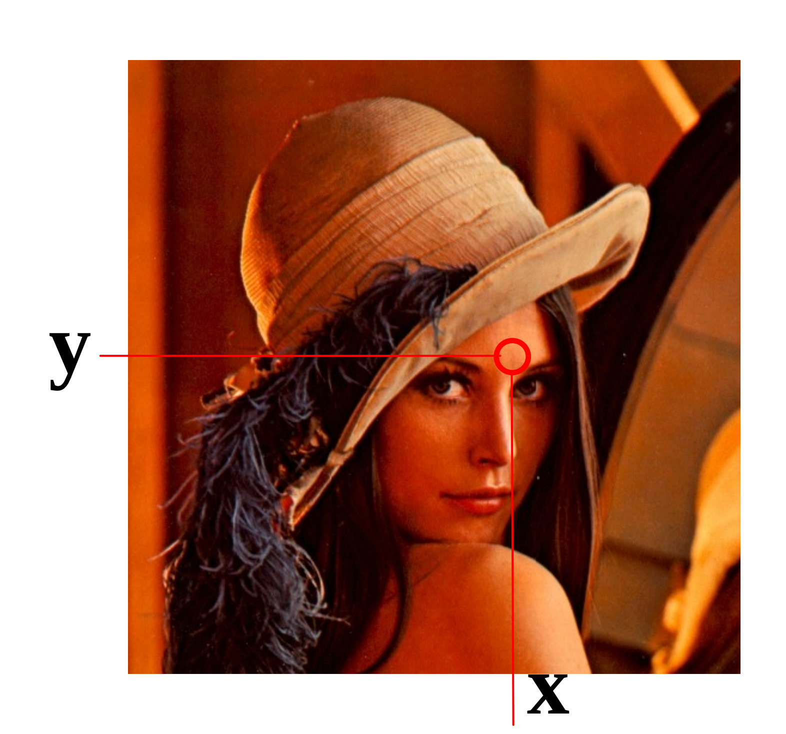
数字图像的像素表示



什么是像素?

数字图像由二维的元素组成，每一个元素具有一个特定的位置（x, y）和幅值f(x, y), 这些元素就称为像素







数字图像处理的起源

- 人类分析



✓ 空间应用：图像增强、图像复原

✓ 医学图像……

✓ 作用：通过图像增强、图像复原等技术，增强或复原模糊或损毁的图像

- 机器感知

✓ 自动字符识别OCR

✓ 人脸识别

✓ 指纹识别，生物特征识别……





- **传统领域**
  - 医学、空间应用、地理学、生物学、军事……
- **最新领域**
  - 数码相机(DC)、数码摄像机(DV)
  - 指纹识别、人脸识别
  - 互联网、视频、多媒体等
  - 基于内容的图像检索、视频检索、多媒体检索
  - 水印、游戏、电影特技、虚拟现实、电子商务等

数字图像处理的应用无处不在

例1：图像增强——直方图均衡化



例2：人脸检测与识别



例3：镜头边界检测





例4：基于内容的图像检索

例5：基于内容的镜头检索



方正天资媒体资产管理系统

视频检索　操作(O)　查看(V)　帮助(H)

选择视频文件　G:\新HSV特征　显　示　提取特征　检　索

Static

022430.bmp 022542.bmp 022588.bmp 022663.bmp 022704.bmp 022756.bmp 022825.bmp 022930.bmp 022965.bmp 023047.bm

Static

022430.bmp 026305.bmp 025784.bmp 025973.bmp 024696.bmp 026698.bmp 024078.bmp 025458.bmp 022930.bmp 022756.bm

026494.bmp 023047.bmp 023252.bmp 022663.bmp 026255.bmp 023172.bmp 024561.bmp 031076.bmp 016580.bmp 025861.bm

023985.bmp 015871.bmp 023838.bmp 024384.bmp 025070.bmp 023633.bmp 015837.bmp 026201.bmp 025130.bmp 022965.bm

019686.bmp 034182.bmp 023703.bmp 025337.bmp 026098.bmp 022704.bmp 034936.bmp 024032.bmp 026782.bmp 020440.bm

015851.bmp 019901.bmp 023786.bmp 023690.bmp 034397.bmp 034402.bmp 030633.bmp 026395.bmp 016079.bmp 022588.bm

例6：基于内容的视频片断检索



方正天资媒体资产管理系统

相似性检索　精确检索　帮助

059064.bmp　059079.bmp　059099.bmp　059108.bmp　059124.bmp　059175.bmp　059231.bmp　059285.bmp　059333.

059064.bmp　059079.bmp　059099.bmp　059108.bmp　059124.bmp　059175.bmp　059231.bmp　059285.bmp　059333.bm

090773.bmp　090793.bmp　090816.bmp　090835.bmp　090876.bmp　090906.bmp　090918.bmp　090927.bmp　090966.bm

226667.bmp　226683.bmp　226697.bmp　226750.bmp　226806.bmp　226817.bmp　226847.bmp

137049.bmp　137098.bmp　137151.bmp　137163.bmp　137169.bmp　137221.bmp　137229.bmp　137247.bmp　137267.bm

例7：视频字幕识别



MpegToImage

控制

OpenFile　Play

GetImage　Exit

字幕区域提取模块

1

2

二值化模块

1　2

3　4

背景去噪模块

1

2

同步方式

1

2

CCTV 1

我又不是在这儿闲着呢

票贩
商业企业的
和餐饮的
最大额度能开到多少
能开90多万
一张3十
等
他赵我谈
你就不要过来
我又不畏导你在迹
我是过路的
到前面去转转
我又不是在这儿闲着呢

例7：视频字幕识别



MpegToImage

控制

OpenFile　Play

GetImage　Exit

字幕区域提取模块

1

2

二值化模块

1　2

3　4

背景去噪模块

1

2

同步方式

1

2

对了，送她一盒糖果，10块钱，你知道的那种

00003944

我们闹别扭
问你太太好
我的建议不变
尽量在控制Bit间段内散布好消息
让对手获得高度评价吧，我们只要去银行诉苦
“周我们一块吃午饭
萨姆，期待你的来信，祝你快乐·
萨姆7◇待你的来信，祝你快乐·
咱们走到大楼去
　好消耗血糖，孩子，快点
我没有穿外套
格雷琴　萨宾森
对了送她一盒糖果，10块钱，你知道的那种



例8：101个视频概念的检测

Aircraft L. Allawi Anchor Animal Y. Arafat Baseball Basketball Beach Bicycle Bird T. Blair Boat

Building Bus G. Bush jr. G. Bush sr. Candle Car Cartoon Chair Charts B. Clinton Cloud Corp. Leader

Court Crowd Cycling Desert Dog Drawing Drawing & Cartoon Duo-anchor Entertainment Explosion Face Female

Fire weapon Fish Flag Flag USA Food Football Golf Government building Government leader Graphics Grass Horse

Horse racing House Indoor H. Jintao J. Kerry E. Lahoud Male Map Meeting Military Monologue Motorbike

Mountain H. Nasrallah Natural disaster News paper Night fire Office Outdoor Overlayed text People People marching People walking Police/ security

C. Powell Prisoner Racing Religious leader River Road Screen A. Sharon Sky Smoke Snow Soccer

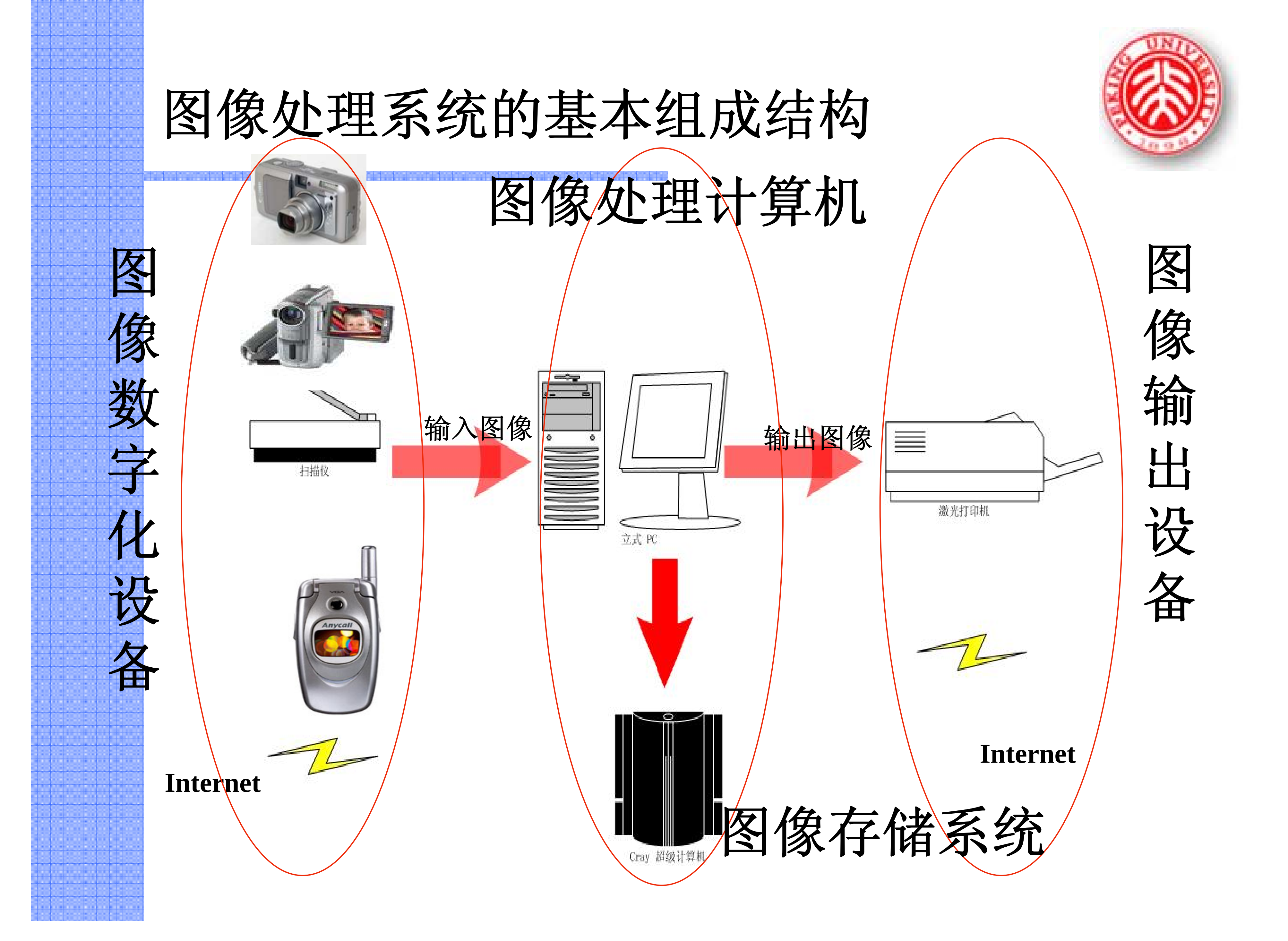
Split screen Sports Studio Swimming pool Table Tank Tennis Tower Tree Truck Urban Vegetation

Vehicle Violence Waterfall Waterscape Weather

例8：101个视频概念的检测



视频语义标注的基本方法


图像处理系统的基本组成结构
图像处理计算机
图像数字化设备
扫描仪
Internet
输入图像
立式 PC
输出图像
激光打印机
图像输出设备
Internet
Cray 超级计算机
图像存储系统

图像处理系统的基本组成结构



- 主要由三大部分组成
  - ✓ 图像数字化设备，包括数码相机、数码摄像机、带照相和/或摄像功能的手机等
  - ✓ 图像处理设备，包括计算机和存储系统
  - ✓ 图像输出设备，包括打印机，也可以输出到Internet上的其它设备

图像存储系统

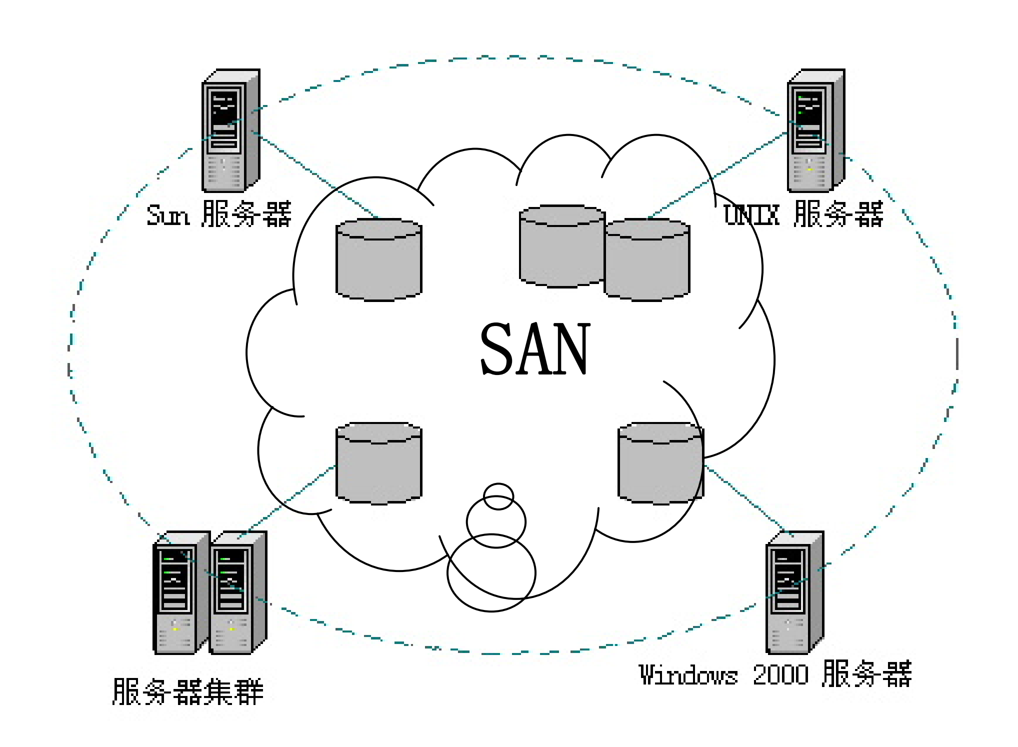


- 图像文件格式体系
  - 互联网用：GIF、JPG
  - 印 刷 用：TIF、JPG、TAG 、PCX
  - 国际标准：TIF、JPG、BMP
- 图像存储体系：分级存储
  - 内存存储：处理时使用
  - 硬盘存储：处理、备份时用（在线）
  - 备份存储：光盘、磁带（离线、近线）
  - 网络存储：SAN、 NAS

为什么要用SAN



存储区域网SAN(Storage Area Network)



SAN是什么？



Client
Client
Server
Router
WAN
LAN
Server
Client
Client
Server
HUB
Switch
Switch
Disk array
Hub
Disk array
Disk array
Switch
Switch
Tape Libarary
Disk array
Disk array
Disk array
Tape Libarary



SAN是什么？（续）

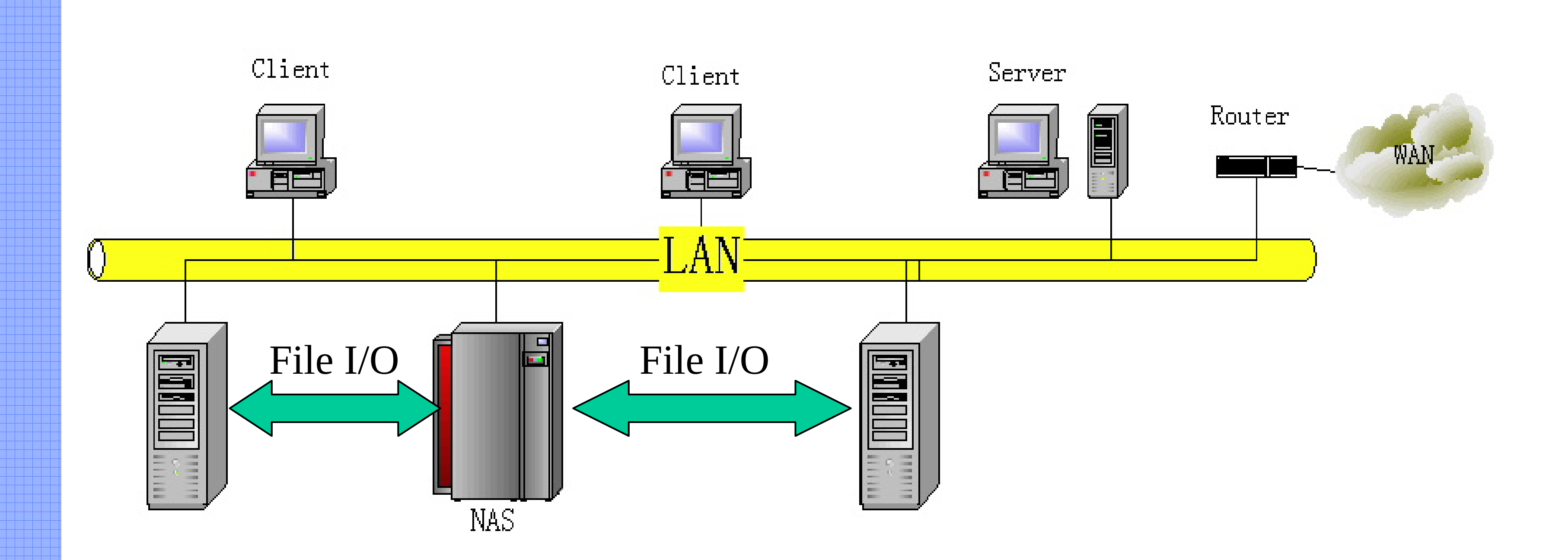


- 不是client/server,而是client/storage devices
- 独立于LAN之外的高速存储网络
- 一般采用高速的光纤通道作为传输媒体（2Gbit/s）
- 将存储设备通过光通道互连设备构成一个存储子网
- 支持服务器和存储设备之间任意到任意的连接
- SAN上的任何一台服务器均可存取网络中的任何一个存储设备
- 对网上的存储资源实施集中统一的管理

NAS (Network Attached Storage)是什么？



是优化的文件服务器，存储设备与服务器均直接连接到LAN上，使用TCP/IP等LAN协议，通过LAN实现数据交换和存储管理。由于使用网络协议，因而会有速度和延时的问题，且系统扩展能力受到网络带宽的限制。



SAN与NAS比较



SAN	NAS
块级共享	文件级共享
远程存储访问	远程文件访问
存储专用网	共享LAN
存储协议（如FCP)	网络协议(如TCP/IP)
集中式管理	分散式管理
无限的扩展能力	有限的扩展能力
更高的连接速度和处理能力	较低的连接速度和处理能力

数字图像处理基础



- 图像的采样和量化
- 数字图像的表示
- 数字图像的质量
- 像素间的一些基本关系

邻域处理方法是图像增强和复原过程的核心

图像的采样和量化

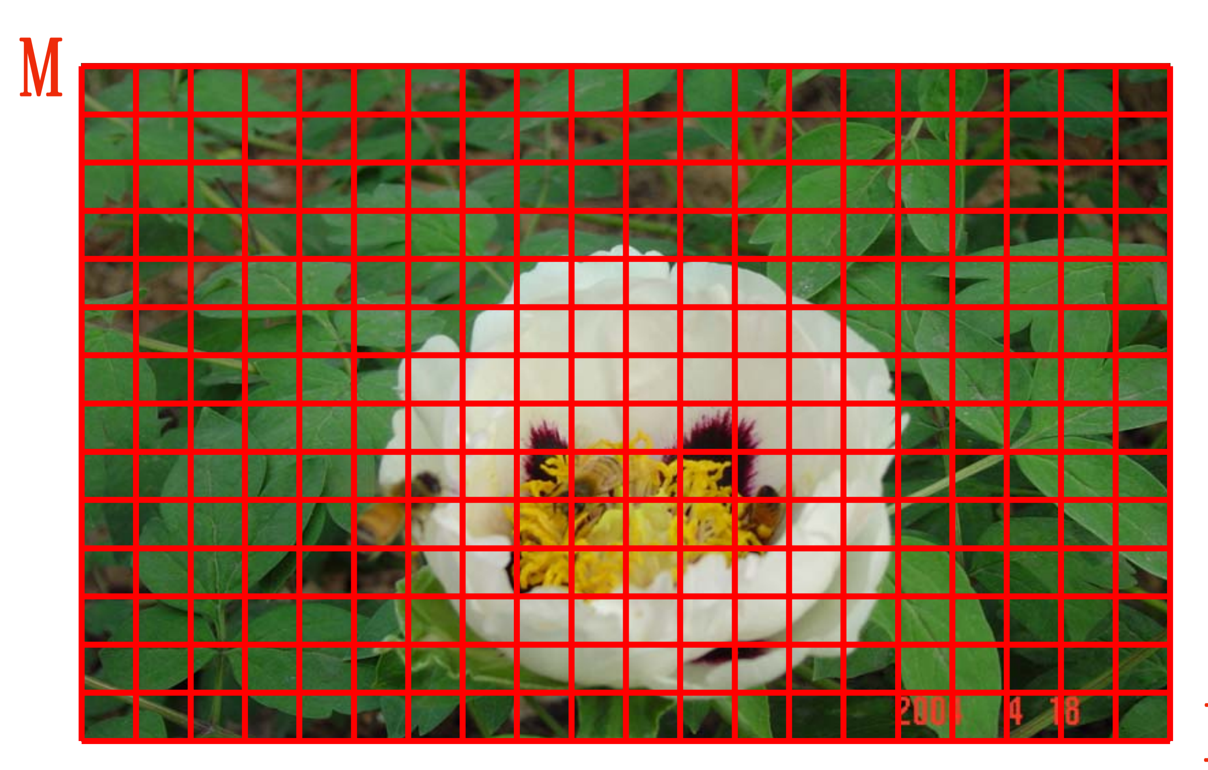


- 大多数传感器的输出是连续电压波形
- 为了产生一幅数字图像，需要把连续的感知数据转化为数字形式
- 这包括两种处理：取样和量化
- 取样：图像空间坐标的数字化
- 量化：图像函数值（灰度值）的数字化





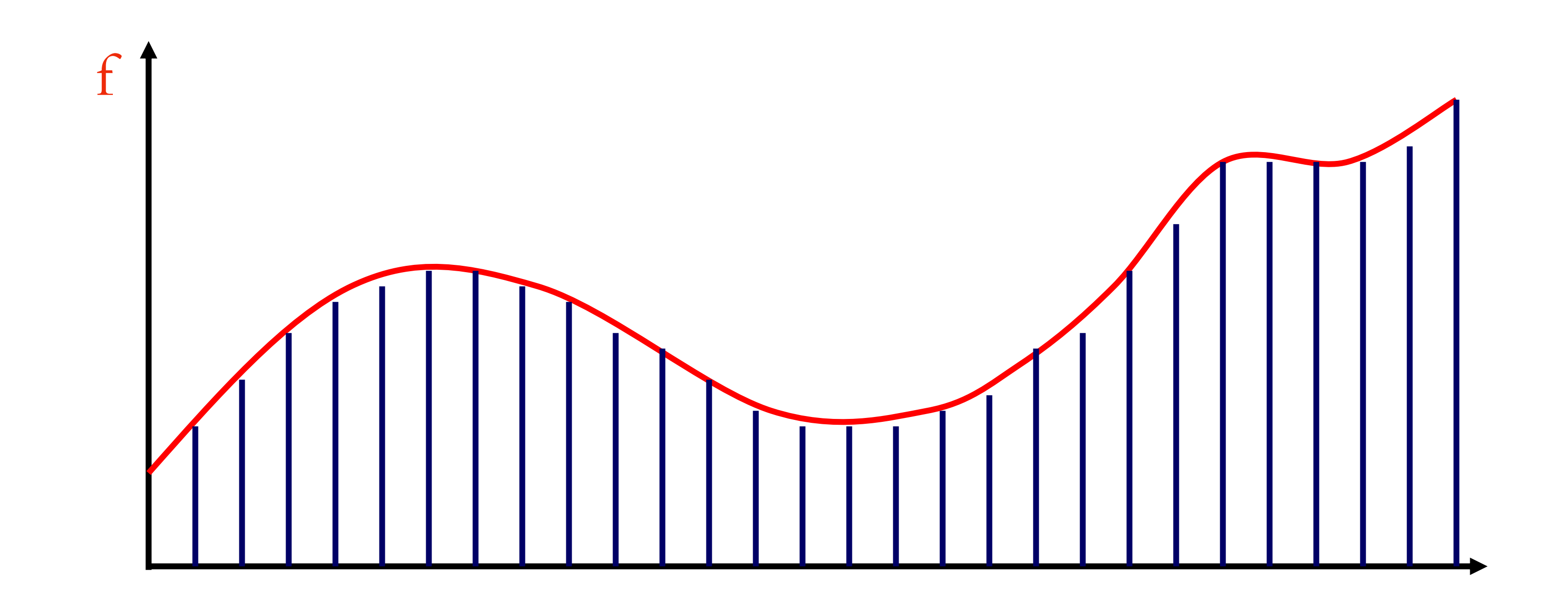
- 空间坐标(x, y)的数字化被称为图像采样
- 确定水平和垂直方向上的像素个数N、M



图像的量化



- 函数取值的数字化被称为图像的量化，如量化到256个灰度级



图像的采样与数字图像的质量



1024 512 256 128 64 32

FIGURE 2.19 A 1024 × 1024, 8-bit image subsampled down to size 32 × 32 pixels. The number of allowable gray levels was kept at 256.

图像的采样与数字图像的质量



1024×1024　　512×512　　256×256

a b c
d e f

128×128　　64×64　　32×32

FIGURE 2.20 (a) 1024 × 1024, 8-bit image. (b) 512 × 512 image resampled into 1024 × 1024 pixels by row and column duplication. (c) through (f) 256 × 256, 128 × 128, 64 × 64, and 32 × 32 images resampled into 1024 × 1024 pixels.

图像的采样与数字图像的质量





265x180



133x90



66x45



33x22

图像的量化与数字图像的质量





256灰度级



16灰度级



8灰度级



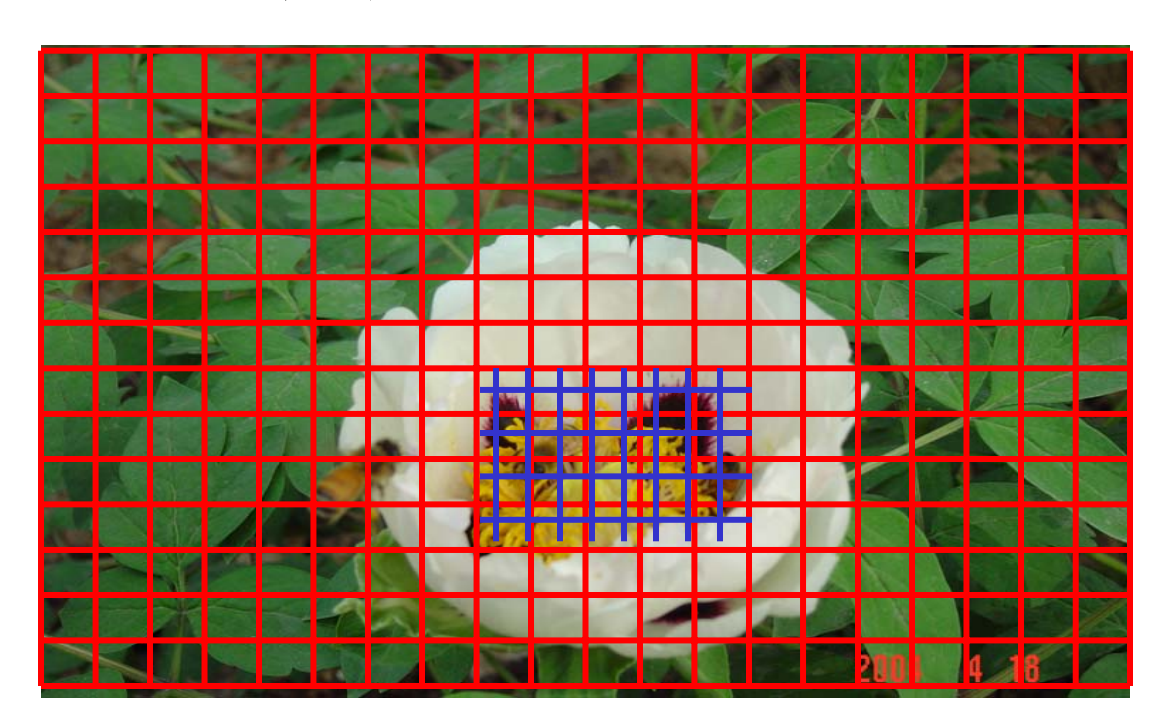
4灰度级



图像的采样和量化



- 非统一的图像的采样
  - ✓ 在灰度级变化尖锐的区域，用细腻的采样，在灰度级比较平滑的区域，用粗糙的采样







- 非统一的图像的量化
  - ✓ 在边界附近使用较少的灰度级。剩余的灰度级可用于灰度级变化比较平滑的区域
  - ✓ 避免或减少由于量化的太粗糙，在灰度级变化比较平滑的区域出现假轮廓的现象

数字图像处理基础



- 图像的采样和量化
- 数字图像的表示
- 数字图像的质量
- 像素间的一些基本关系



数字图像的表示



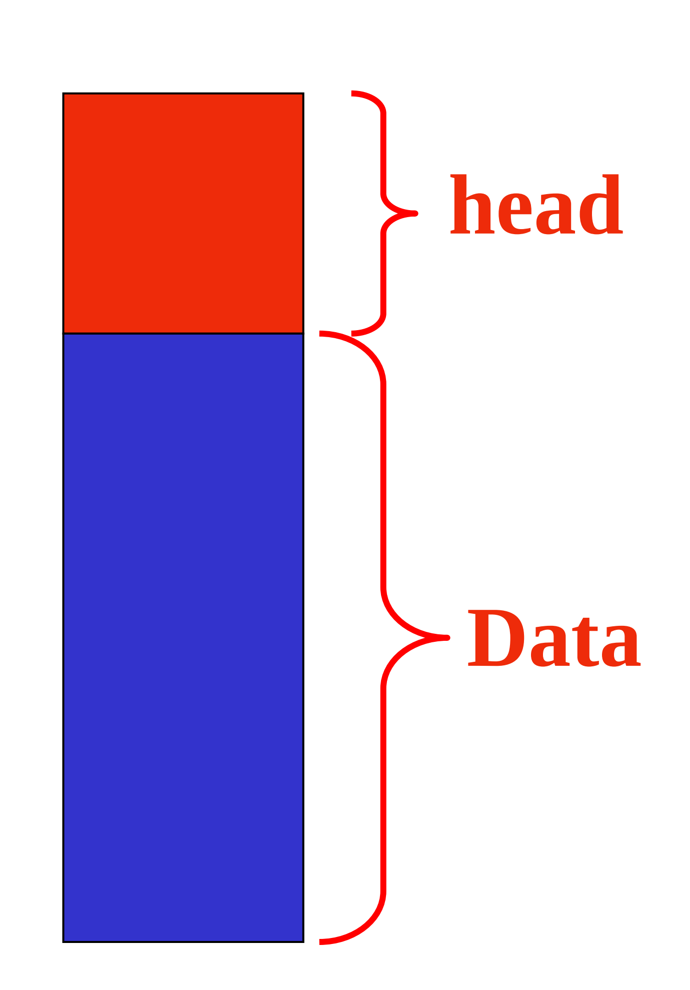
- 二维离散亮度函数——f(x, y)
  - ✓ x，y说明图像像素的空间坐标
  - ✓ 函数值 f 代表了在点(x, y)处像素的灰度值
- 二维矩阵——A[m, n]
  - ✓ m ， n说明图像的宽和高。
  - ✓ 矩阵元素a(i， j)的值，表示图像在第 i 行，第 j 列的像素的灰度值；i, j表示几何位置



数字图像的表示



- 图像描述信息
  - ✓ 如图像高度和宽度等信息
- 图像数据
  - ✓ 顺序存放的连续数据
- BMP格式
  1. 位图文件头
  2. 位图信息头
  3. 调色板
  4. 图像数据



BMP格式——1、位图文件头



```
typedef struct tagBITMAPFILEHEADER{
   WORD    bfType;            //文件类型，必须是字符串"BM"
   DWORD   bfSize;            //指定文件大小
   WORD    bfReserved1;       //保留字，不考虑
   WORD    bfReserved2;       //保留字，不考虑
   DWORD   bfOffBits;         //从文件头到位图数据的偏移字节数
} BITMAPFILEHEADER;
```

BMP格式——2、位图信息头



```
typedef struct tagBITMAPINFOHEADER{
   DWORD  biSize;           //该结构的长度，40个字节
   LONG   biWidth;          //图像的宽度，单位是像素
   LONG   biHeight;         //图像的高度，单位是像素
   WORD   biPlanes;         //必须是1
   WORD   biBitCount        //颜色位数，如1，4，8，24
   DWORD  biCompression;    //压缩类型，如BI_RGB,BI_RLE4
   DWORD  biSizeImage;      //实际位图数据占用的字节数
   LONG   biXPelsPerMeter;  //水平分辨率
   LONG   biYPelsPerMeter;  //垂直分辨率
   DWORD  biClrUsed;        //实际使用的颜色数
   DWORD  biClrImportant;   //重要的颜色数
} BITMAPINFOHEADER;
```



BMP格式——3、调色板



```
typedef struct tagRGBQUAD{
    BYTE    rgbBlue;              //该颜色的蓝色分量
    BYTE    rgbGreen;             //该颜色的绿色分量
    BYTE    rgbRed;               //该颜色的红色分量
    BYTE    rgbReserved;          //保留值，不考虑
} RGBQUAD;
```

注：有些位图不需要调色板，如真彩色图，它们的BITMAPINFOHEADER后面直接是位图数据

BMP格式——4、实际的图像数据



- 对于2色位图，1位表示一个像素颜色，所以一个字节表示8个像素
- 对于16色位图，4位表示一个像素颜色，所以一个字节表示2个像素
- 对于256色位图，1个字节表示1个像素
- 对于真彩色图，3个字节表示一个像素

数字图像处理基础



- 图像的采样和量化
- 数字图像的表示
- 数字图像的质量
- 像素间的一些基本关系

图像的质量：1、层次



- 灰度级：表示像素明暗程度的整数量

  例如：像素的取值范围为0-255，就称该图像为256个灰度级的图像

- 层　　次：表示图像实际拥有的灰度级的数量

  例如：具有32种不同取值的图像，可称该图像具有32个层次

图像数据的实际层次越多，视觉效果就越好

图像的质量：1、层次



256个层次的图像

64个层次的图像

16个层次的图像

图像的质量：2、对比度



- 对比度：是指一幅图像中灰度反差的大小

  对比度 = 最大亮度 / 最小亮度





图像的质量：3、清晰度



- 与清晰度相关的主要因素
  - ✓亮度
  - ✓对比度
  - ✓尺寸大小
  - ✓细微层次
  - ✓颜色饱和度

影响清晰度因素——1、亮度



原图



降低亮度



影响清晰度因素——2、对比度



原图



降低对比度



影响清晰度因素——3、尺寸大小



原图







影响清晰度因素——4、细微层次



原图



减少细微层次



影响清晰度因素——5、颜色饱和度



原图

降低颜色饱和度





数字图像处理基础



- 图像的采样和量化
- 数字图像的表示
- 数字图像的质量
- 像素间的一些基本关系

像素间的一些基本关系

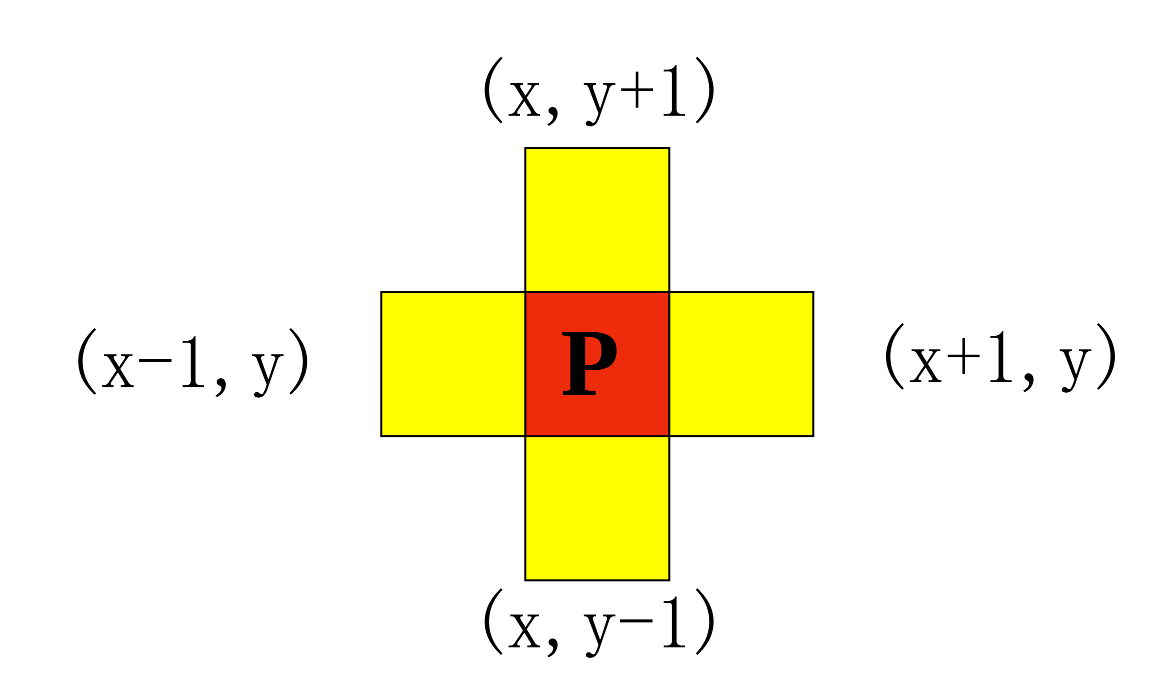


- 相邻像素：
  - ✓ 4邻域
  - ✓ D邻域
  - ✓ 8邻域
- 连通性
  - ✓ 4连通
  - ✓ 8连通
  - ✓ m连通
- 距离

相邻像素——4邻域



- 4邻域：像素p(x, y)的4邻域是：
  (x+1, y)；(x−1, y)；(x, y+1)；(x, y−1)
- 用$N_4(p)$表示像素p的4邻域



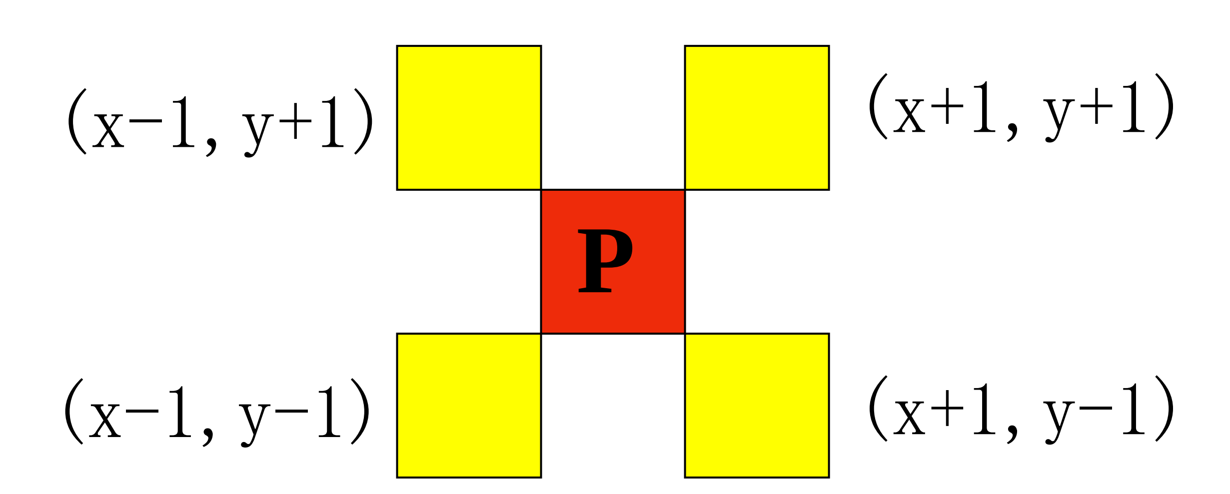
相邻像素——D邻域



- D邻域定义：像素p(x, y)的D邻域是：

  对角上的点 (x+1, y+1)；(x+1, y-1)；(x-1, y+1)；(x-1, y-1)

- 用$N_D(p)$表示像素p的D邻域



相邻像素——8邻域

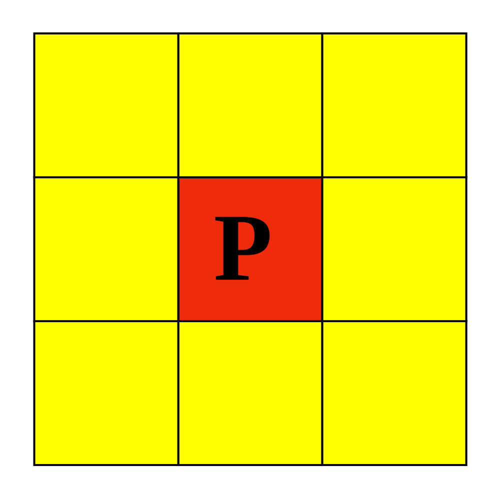


- 8邻域定义：像素p(x, y)的8邻域是：

  4邻域的点 ＋ D邻域的点

- 用$N_8(p)$表示像素p的8邻域。

$$N_8(p) = N_4(p) + N_D(p)$$



像素间的连通性

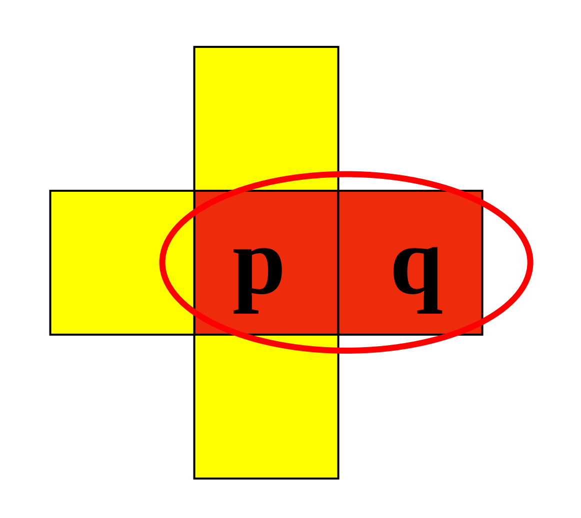


- 连通性是描述区域和边界的重要概念
- 两个像素连通的两个必要条件是：
  - ✓ 两个像素的位置是否相邻
  - ✓ 两个像素的灰度值是否满足特定的相似性准则（或者是否相等）
- 4连通、8连通、m连通的定义

像素的连通性——4连通



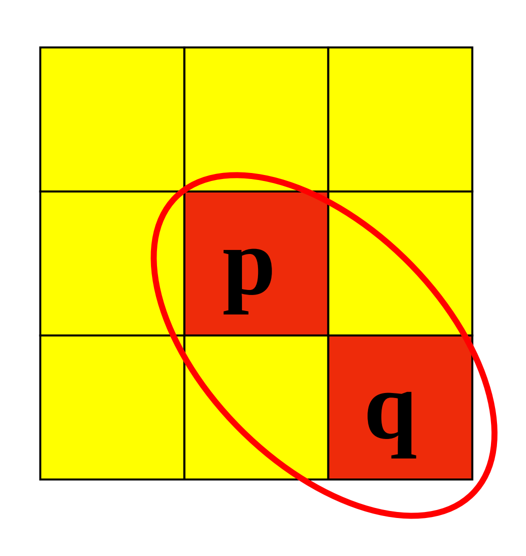
- 对于具有值V的像素p和q，如果q在集合$N_4(p)$中，则称这两个像素是4连通的



像素的连通性——8连通



- 对于具有值V的像素p和q，如果q在集合$N_8(p)$中，则称这两个像素是8连通的



像素的连通性——m连通



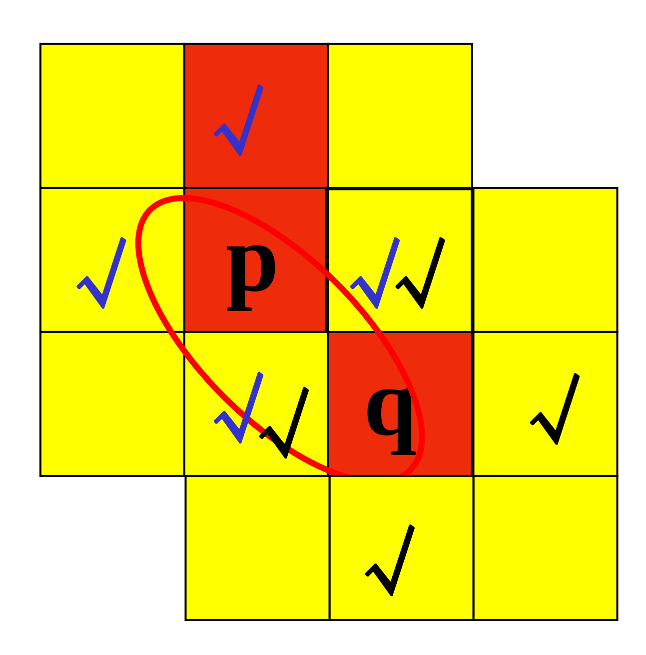
- 对于具有值V的像素p和q，如果：
  I. q在集合$N_4(p)$中，或
  II. q在集合$N_D(p)$中，并且$N_4(p)$与$N_4(q)$的交集为空（没有值V的像素）

则称这两个像素是m连通的，即4连通和D连通的混合连通。

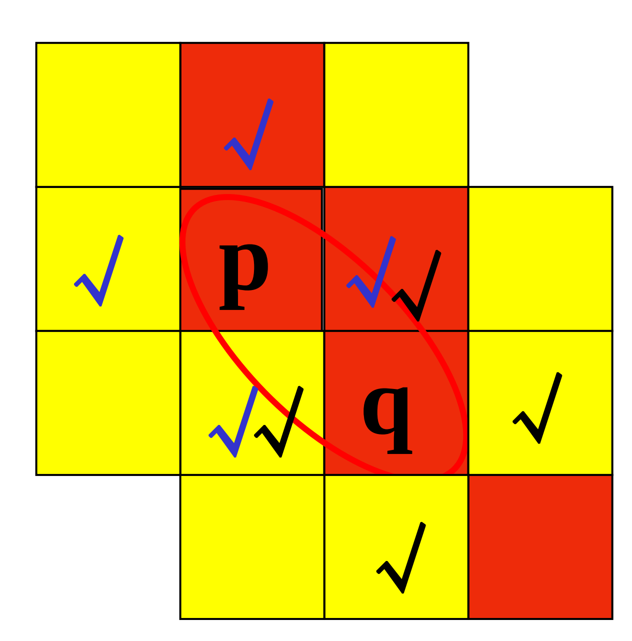
像素的连通性—— m连通



是m连通



不是m连通



像素的连通性——通路



- 通路的定义

  一条从具有坐标(x, y)的像素p, 到具有坐标(s, t)的像素q的通路，是具有坐标

  $(x_0, y_0), (x_1, y_1), . . . , (x_n, y_n)$的不同像素的序列。其中，$(x_0, y_0) = (x, y)$，$(x_n, y_n) = (s, t)$，$(x_i, y_i)$ 和(x_{i-1}, y_{i-1})是邻接的，$1 \leqslant i \leqslant n$，n是路径的长度。如果$(x_0, y_0) = (x_n, y_n)$，则该通路是闭合通路

像素的连通性——距离



- 像素之间距离的定义
- 欧氏距离定义
- D_4距离（城市距离）定义
- D_8距离（棋盘距离）定义

像素之间距离的定义



对于像素p、q和z，分别具有坐标(x, y)，(s, t)和(u, v)，如果

(1) $D(p, q) \geq 0$ （$D(p, q)=0$，当且仅当$p=q$），

(2) $D(p, q) = D(q, p)$

(3) $D(p, z) \leq D(p, q) + D(q, z)$

则称D是距离函数或度量

欧式距离定义



- 像素p(x, y)和q(s, t)间的欧式距离定义如下：

$$D_e(p,q)=\sqrt{(x-s)^2+(y-t)^2}$$

- 对于这个距离计算法，具有与(x, y)距离小于等于某个值r的像素是：包含在以(x, y)为圆心，以r为半径的圆平面

D_4距离（城市距离）



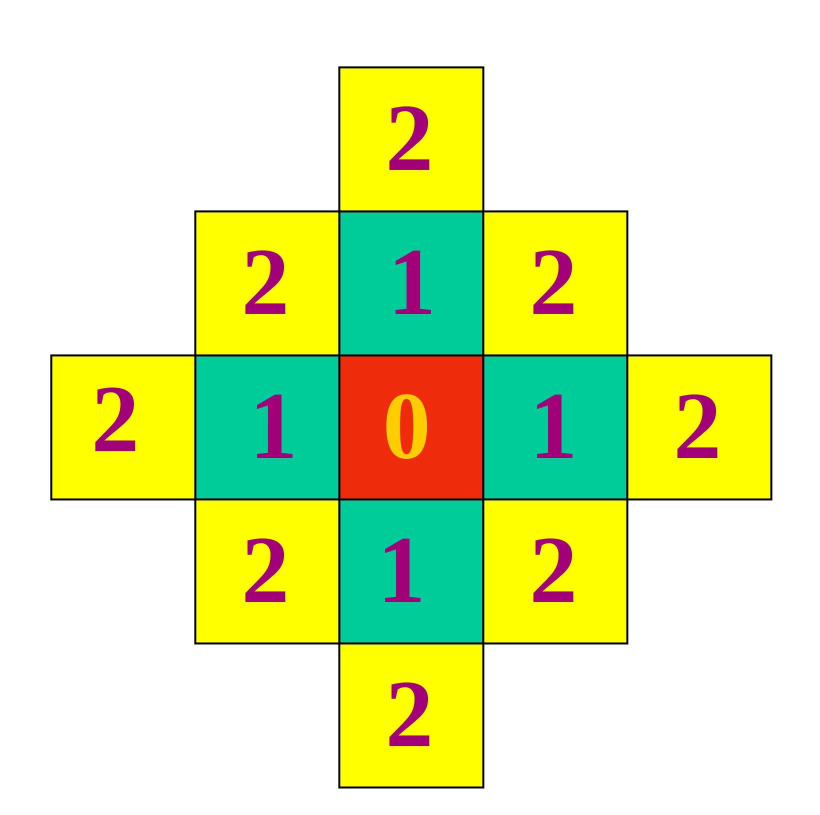
像素p(x, y)和q(s, t)之间的D_4距离定义为：

$$D_4(p, q) = |x - s| + |y - t|$$

D_4距离举例



- 具有与(x, y)距离小于等于某个值r的那些像素形成一个菱形
- 例如，与点(x, y)（中心点）D_4距离小于等于2的像素，形成右边固定距离的轮廓
- 具有D_4 = 1的像素是(x, y)的4邻域



D_8距离（棋盘距离）



像素p(x, y)和q(s, t)之间的D_8距离定义为：

$$D_8(p, q) = \max(|x - s|, |y - t|)$$

D_8距离举例



- 具有与(x, y)距离小于等于某个值r的那些像素形成一个正方形
- 例如，与点(x, y)（中心点）D_8距离小于等于2的像素，形成右边固定距离的轮廓
- 具有D_8 = 1的像素是(x, y)的8邻域

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

数字图像处理（1）



任何问题？

研究生课程



数字图像处理

Digital Image Processing

彭宇新

北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

研究生课程



图像增强
- 处理方法
  - 空域方法
    - 点处理（变换）
    - 模板处理（滤波）
  - 频域方法
- 处理策略
  - 全局处理
  - 局部处理
- 处理对象
  - 灰度图像
  - 彩色图像

空间域图像增强



- 基础知识
  - ✓ 基本概念
  - ✓ 点运算
  - ✓ 代数运算
  - ✓ 直方图运算
  - ✓ 应用实例——镜头边界的检测（补充知识）
- 空间滤波器
  - ✓ 平滑空间滤波器
  - ✓ 锐化空间滤波器

基本概念



- 图像增强分为两类：
  - ✓ 空间域增强：对图像的像素直接处理
  - ✓ 频域增强 ： 修改图像的傅里叶变换（后面介绍）
- 空间域增强： $g(x, y) = T[f(x, y)]$
  - ✓ f(x, y)是原图像
  - ✓ g(x, y)是处理后的图像
  - ✓ T是作用于f的操作，定义在(x, y)的邻域
- 空间域增强的简化形式： $s = T(r)$
  - ✓ r是f(x, y)在任意点(x, y)的灰度级
  - ✓ s是g(x, y)在任意点(x, y)的灰度级

点运算



FIGURE 3.3 Some basic gray-level transformation functions used for image enhancement.

Output gray level, s

$L-1$

$3L/4$

$L/2$

$L/4$

0

Negative

nth root

Log

nth power

Identity

Inverse log

0 $L/4$ $L/2$ $3L/4$ $L-1$

Input gray level, r

点运算——1反转变换，2对数变换

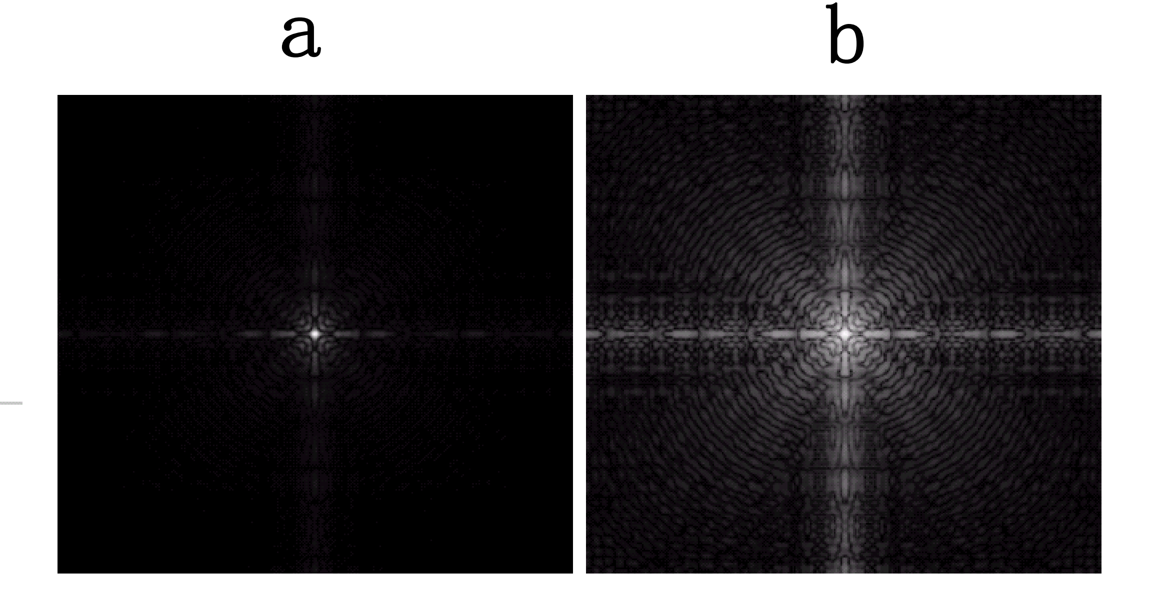


- 反转变换： $s=(L-1)-r$
  - ✓ [0, L−1]为图像的灰度级。作用：黑的变白，白的变黑
- 对数变换： $s=c\log(1+r)$
  - ✓ c是常数
  - ✓ $r \geq 0$
  - ✓ 有时原图的动态范围太大，超出某些显示设备的允许动态范围，如直接使用原图，则一部分细节可能丢失
  - ✓ 解决办法是对原图进行灰度压缩，如对数变换

点运算——1反转变换，2对数变换



- 图a显示了值为0- 1.5×10^6 的傅里叶频谱
- 在一个8位的系统中显示
- 图b显示了对数变换在8位系统中的显示结果



a b

FIGURE 3.5 (a) Fourier spectrum. (b) Result of applying the log transformation given in Eq. (3.2-2) with $c = 1$.



点运算——3幂次变换

- 幂次变换：　$s = cr^{\gamma}$
  - ✓ c和 γ 是正常数
  - ✓ γ <1 提高灰度级，在正比函数上方，使图像变亮
  - ✓ γ >1 降低灰度级，在正比函数下方，使图像变暗

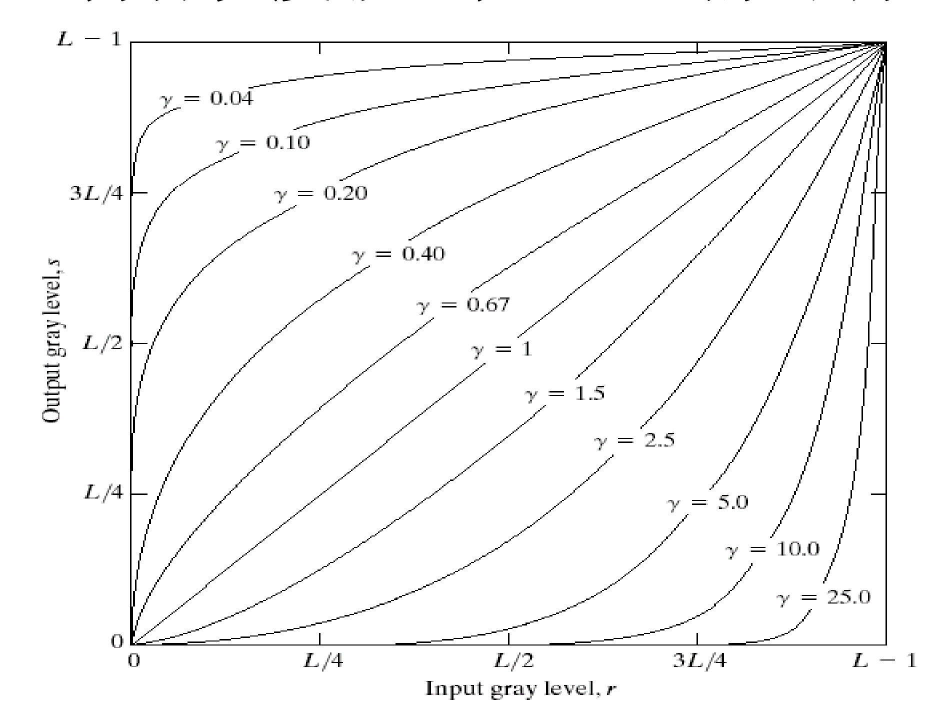


FIGURE 3.6 Plots of the equation $s = cr^{\gamma}$ for various values of γ ($c = 1$ in all cases).



点运算——3幂次变换

- 例：人体胸上部脊椎骨折的核磁共振图像
- $\gamma<1$ 提高灰度级，使图像变亮。c=1, $\gamma=0.6,0.4,0.3$



a b
c d

FIGURE 3.8
(a) Magnetic resonance (MR) image of a fractured human spine.
(b)–(d) Results of applying the transformation in Eq. (3.2-3) with $c = 1$ and $\gamma = 0.6, 0.4,$ and 0.3, respectively. (Original image for this example courtesy of Dr. David R. Pickens, Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center.)

$\gamma=0.4$

增强效果最好





点运算——3幂次变换

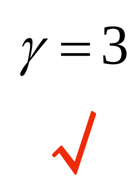
- 例：航空地面图像
- $\gamma > 1$降低灰度级，使图像变暗c=1，$\gamma = 3,4,5$

a	b
c	d

FIGURE 3.9 (a) Aerial image. (b)–(d) Results of applying the transformation in Eq. (3.2-3) with $c = 1$ and $\gamma = 3.0, 4.0$, and 5.0, respectively. (Original image for this example courtesy of NASA.)





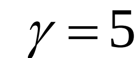


$\gamma = 4$

√





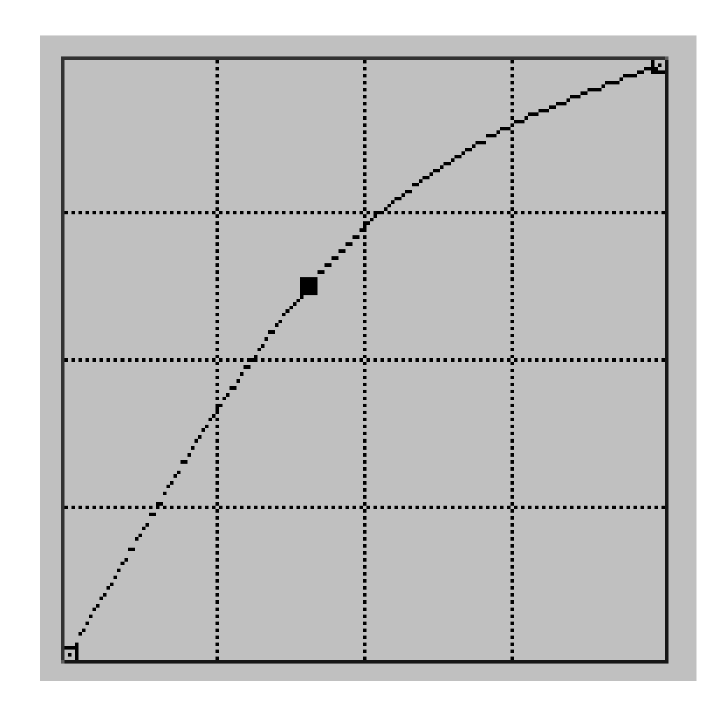


点运算——4对比度拉伸



思想：提高图像处理时灰度级的动态范围







点运算——5灰度级切片



关心范围指定较高值，

其它指定较低值

关心范围指定较高值，

其它保持不变

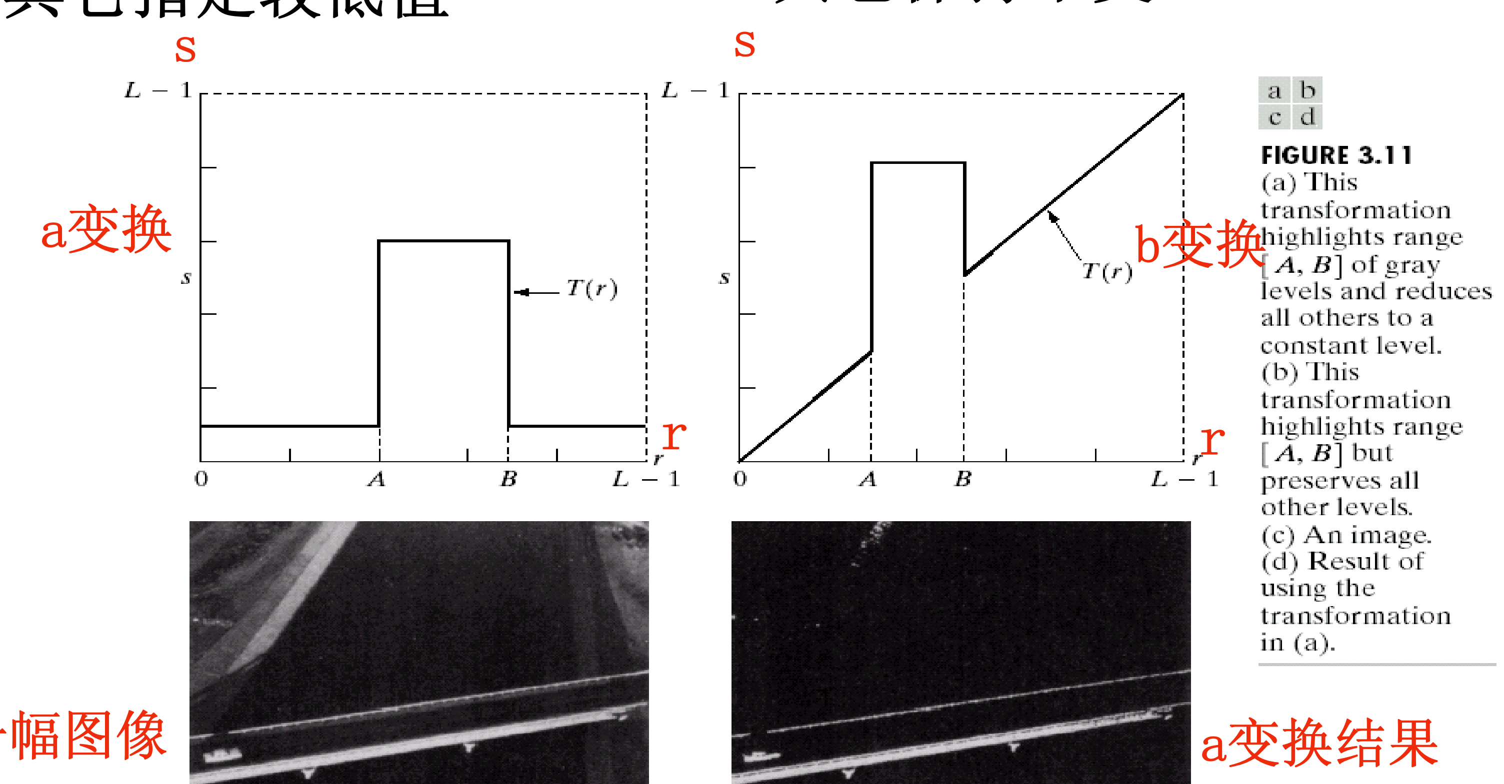


FIGURE 3.11 (a) This transformation highlights range [A, B] of gray levels and reduces all others to a constant level. (b) This transformation highlights range [A, B] but preserves all other levels. (c) An image. (d) Result of using the transformation in (a).

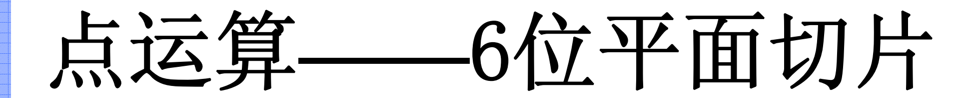
点运算——5灰度级切片



255

0

255



点运算——6位平面切片



- 位平面切片

　假设图像中每个像素的灰度级是256，这可以用8位来表示，假设图像是由8个1位平面组成，范围从位平面0到位平面7。其中，位平面0包含图像中像素的最低位，位平面7包含像素的最高位

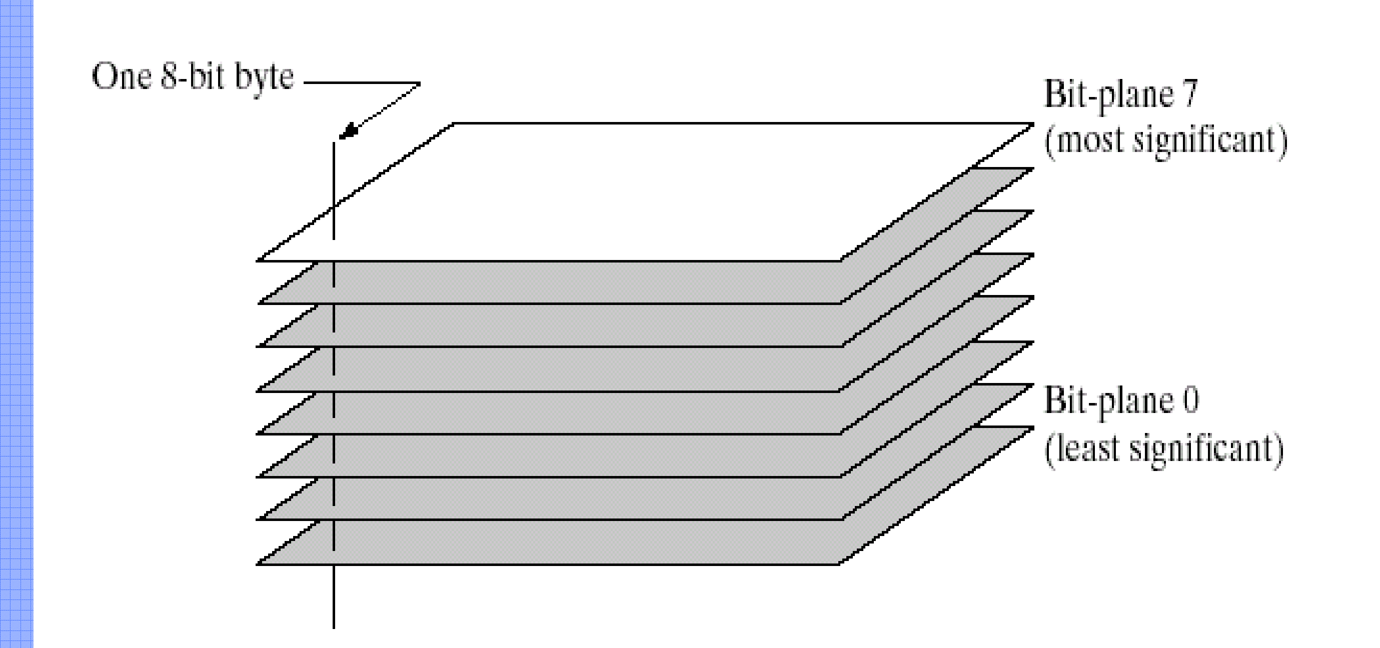


FIGURE 3.12 Bit-plane representation of an 8-bit image.

点运算——6位平面切片



- 作用
  - ✓ 通过对特定位提高亮度，改善图像质量
  - ✓ 较高位（如前4位）包含大多数视觉重要数据
  - ✓ 较低位（如后4位）对图像中的微小细节有作用
  - ✓ 分解为位平面，可以分析每一位在图像中的相对重要性



点运算——6位平面切片

一幅8比特分形图像

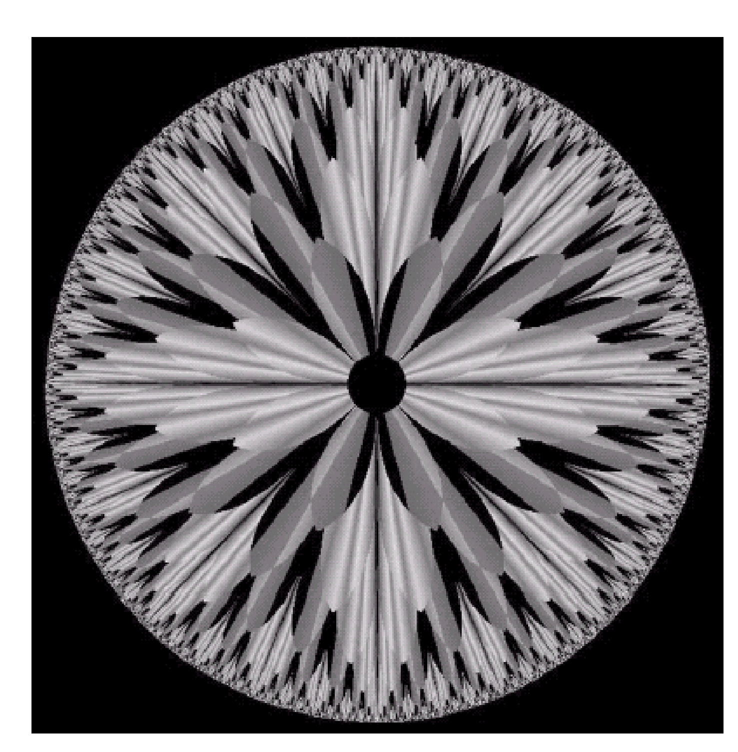


FIGURE 3.13 An 8-bit fractal image. (A fractal is an image generated from mathematical expressions). (Courtesy of Ms. Melissa D. Binde, Swarthmore College, Swarthmore, PA.)

点运算——6位平面切片



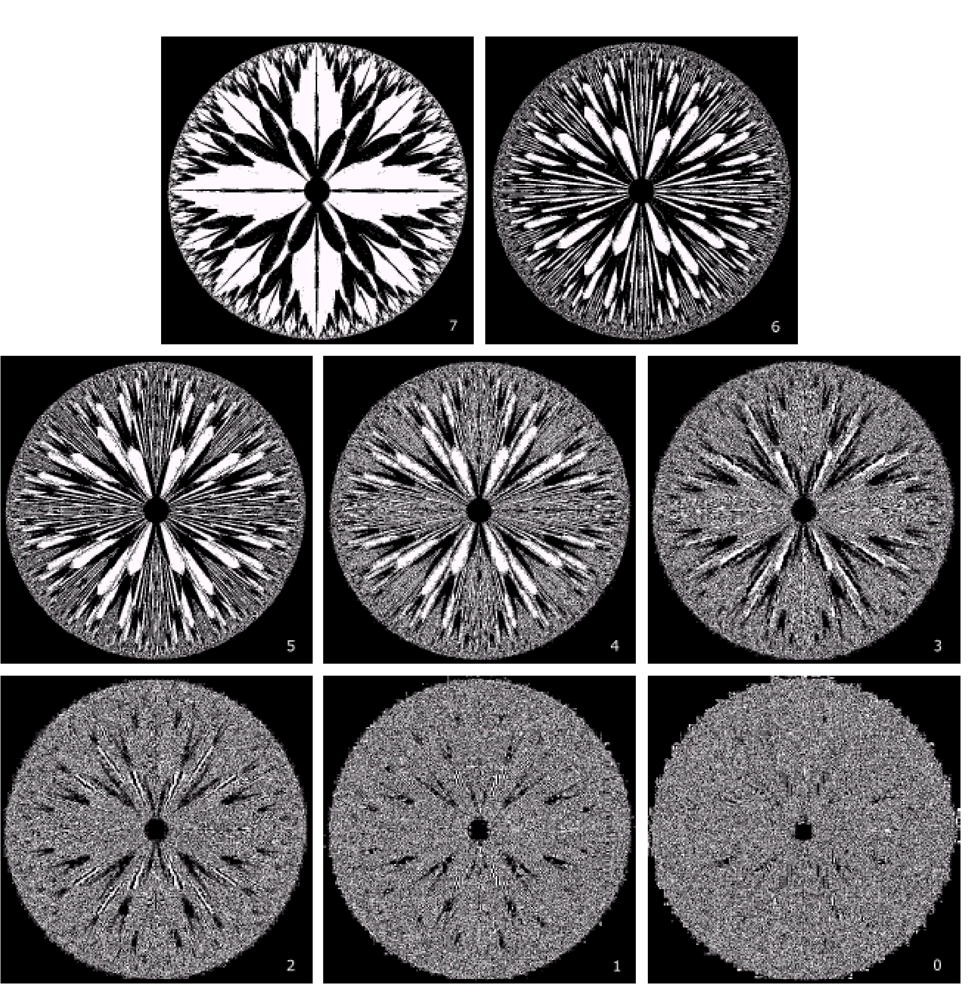


FIGURE 3.14 The eight bit planes of the image in Fig. 3.13. The number at the bottom, right of each image identifies the bit plane.

代数运算



- 算术运算
  - 加
  - 减
  - 乘
  - 除:一幅图像取反和另一幅图像相乘
- 逻辑运算
  - 非
  - 与
  - 或
  - 异或

代数运算——加法



- 加法运算的定义

  $C(x, y) = A(x, y) + B(x, y)$

- 主要应用举例
  - ✓ 去除叠加性噪声
  - ✓ 生成图像叠加效果

代数运算——加法



- 去除叠加性噪声

  对于原图像f(x, y)，有一个噪声图像集

  $\{ g_i(x, y) \} \quad i = 1, 2, ... N$

  其中：$g_i(x, y) = f(x, y) + h(x, y)_i$

  假设噪声h(x, y)均值为0，且互不相关

  N个图像的均值定义为：

  $g(x, y) = 1/N(g_0(x, y)+g_1(x, y)+...+ g_N(x, y))$

  期望值E(g(x, y)) = f(x, y)

  上述图像均值将降低噪声的影响

代数运算——加法



- 去除叠加性噪声——星系图举例

原图

噪声图像

N=8

N=16

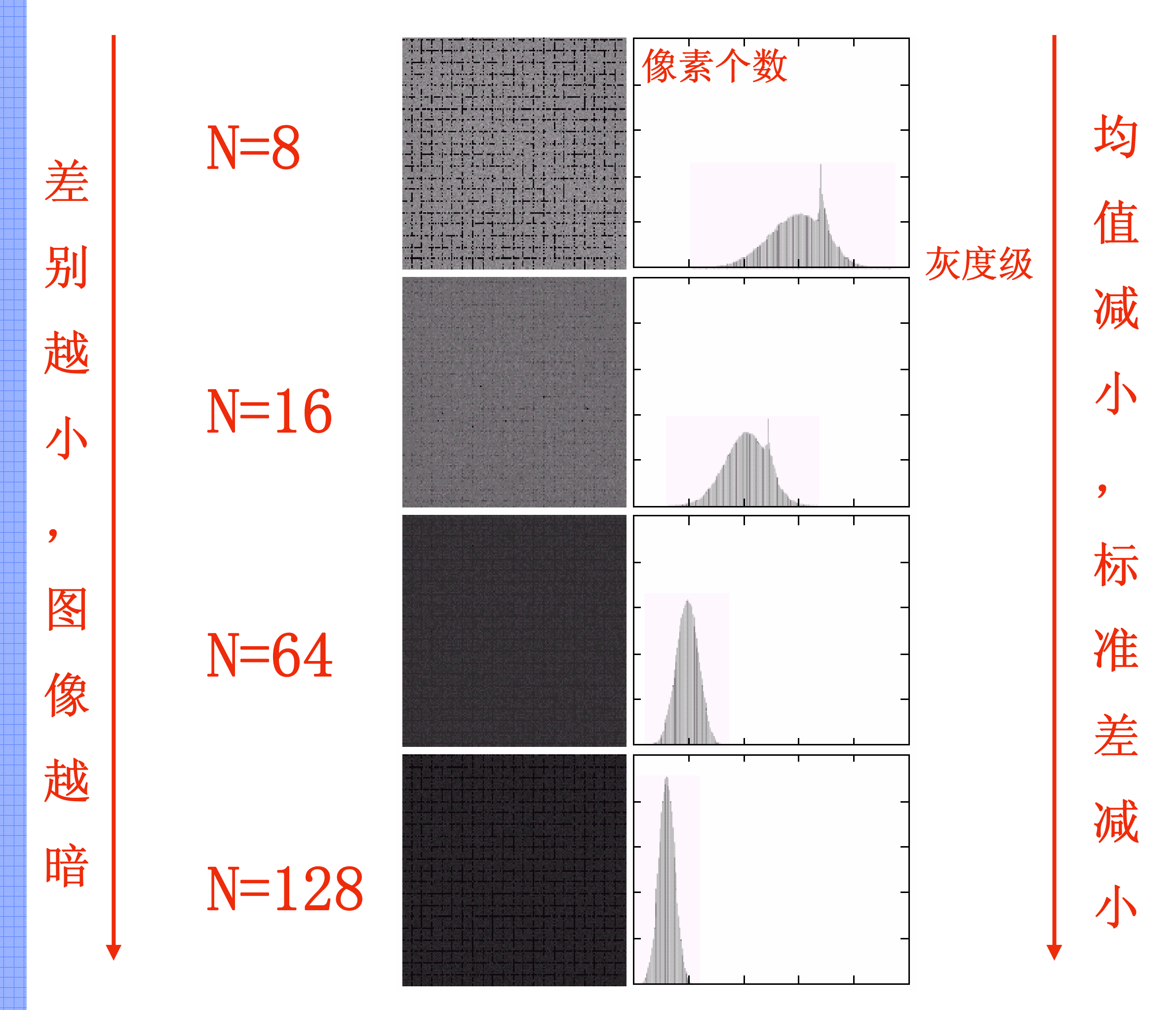
N=64

N=128

去除叠加性噪声——星系图举例



- 原图与均值图像的差值图像和直方图



代数运算——加法



- 生成图像叠加效果

  对于两个图像f(x, y)和h(x, y)的均值有：

  $$g(x,y) = 1/2f(x,y) + 1/2h(x,y)$$

  推广这个公式为：

  $$g(x,y) = \alpha f(x,y) + \beta h(x,y)$$

  其中 $\alpha + \beta = 1$

  可以得到各种图像合成的效果，也可以用于两张图片的衔接

代数运算——加法



代数运算——减法



- 减法的定义

$$C(x,y) = A(x,y) - B(x,y)$$

- 主要应用举例
  - ✓ 显示两幅图像的差异，检测同一场景两幅图像之间的变化

    如：视频中镜头边界的检测
  - ✓ 去除不需要的叠加性图案
  - ✓ 图像分割：如分割运动的车辆，减法去掉静止部分，剩余的是运动元素和噪声

代数运算——减法



- 检测同一场景两幅图像之间的变化

  设：时间1的图像为$T_1(x, y)$，

  时间2的图像为$T_2(x, y)$

$$g(x, y) = T_2(x, y) - T_1(x, y)$$

例：视频中镜头边界的检测(后面介绍)

代数运算——减法



- 去除不需要的叠加性图案

设：背景图像b(x, y)，前景背景混合图像 f(x, y)

$$g(x, y) = f(x, y) - b(x, y)$$

g(x, y) 为去除了背景的图像。

代数运算——减法

- 去除不需要的叠加性图案
  例：电视制作的蓝屏技术

f(x, y)



减去背景b(x, y)

g(x, y)



叠加蓝色背景



问题？

代数运算——乘法



- 乘法的定义

  C(x,y) = A(x,y) * B(x,y)

- 主要应用举例

  ✓ 图像的局部显示

  用二值蒙板图像与原图像做乘法

代数运算——乘法





＝



×



代数运算——1、非



- 非的定义

  $g(x, y) = 255 - f(x, y)$

- 主要应用举例
  - ✓ 获得一个阴图像
  - ✓ 获得一个子图像的补图像



代数运算——1、非

- 获得一个阴图像

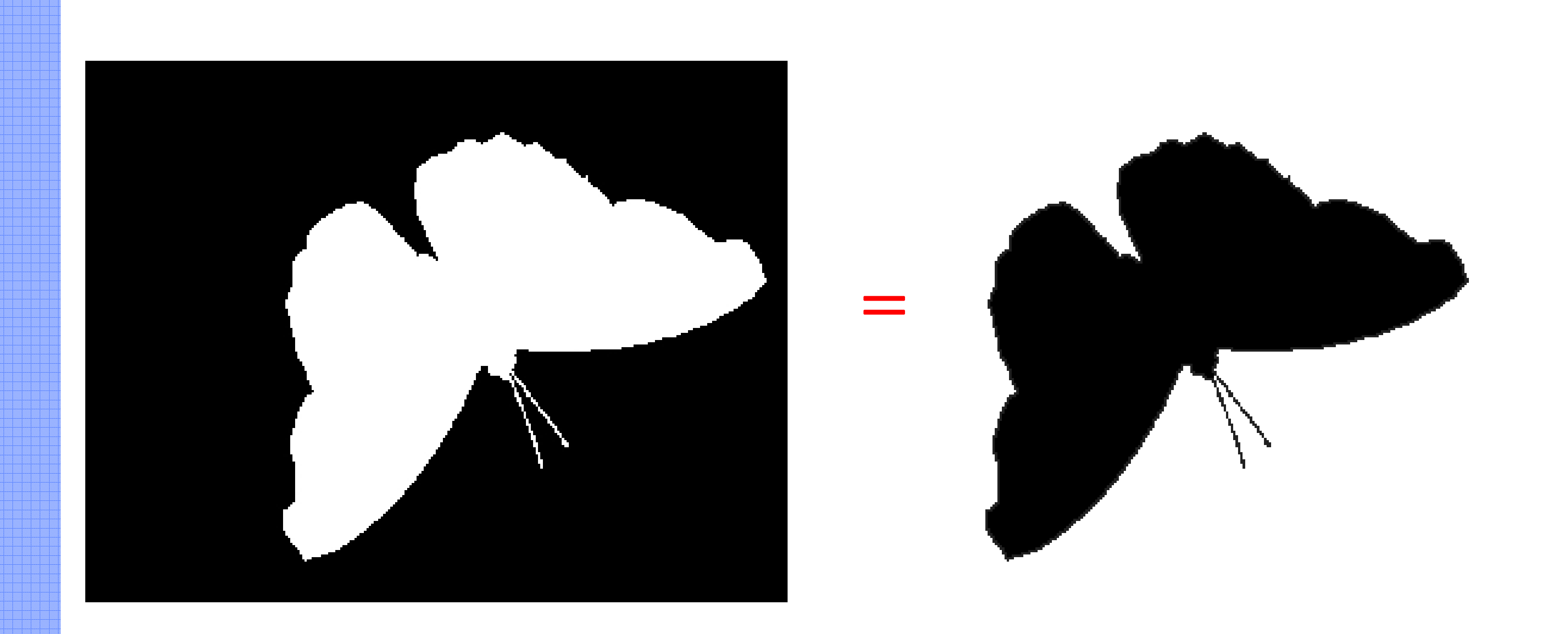




代数运算——1、非



- 获得一个子图像的补图像





代数运算——2、与

- 与运算的定义

$$g(x, y) = f(x, y) \wedge h(x, y)$$

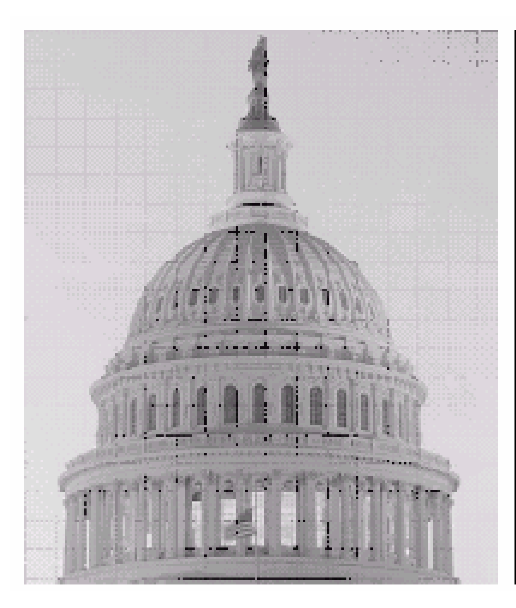
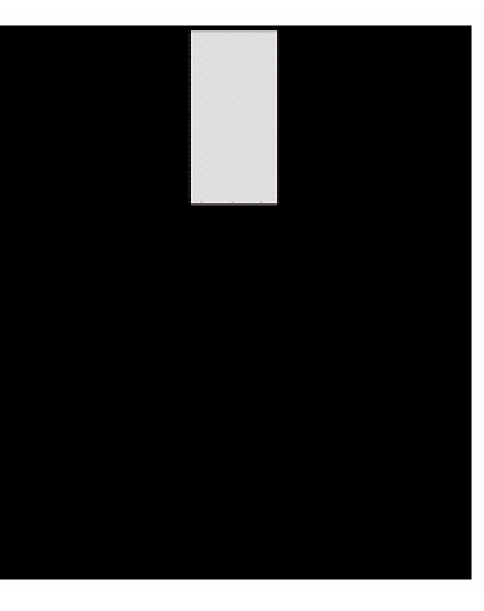
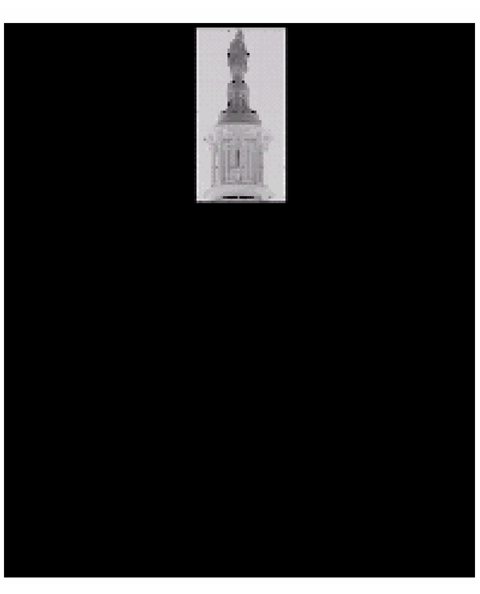
- 主要应用举例
  - ✓ 求两个子图像的相交子图



代数运算——2、与



- ✓ 模板运算：提取感兴趣的子图像

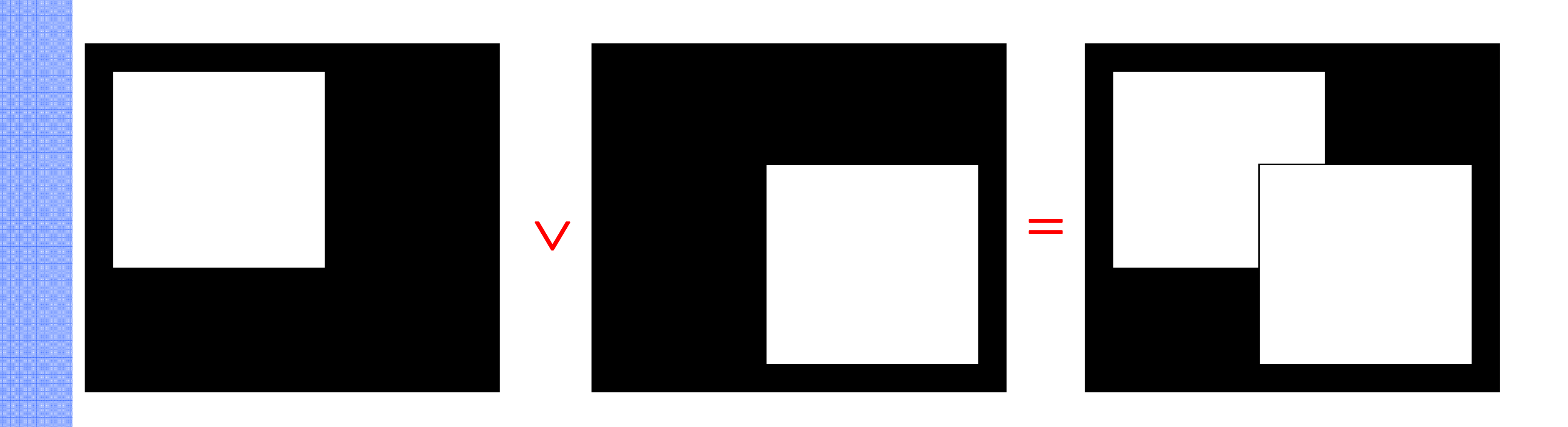
代数运算——3、或



- 或运算的定义

  $g(x,y) = f(x,y) \vee h(x,y)$

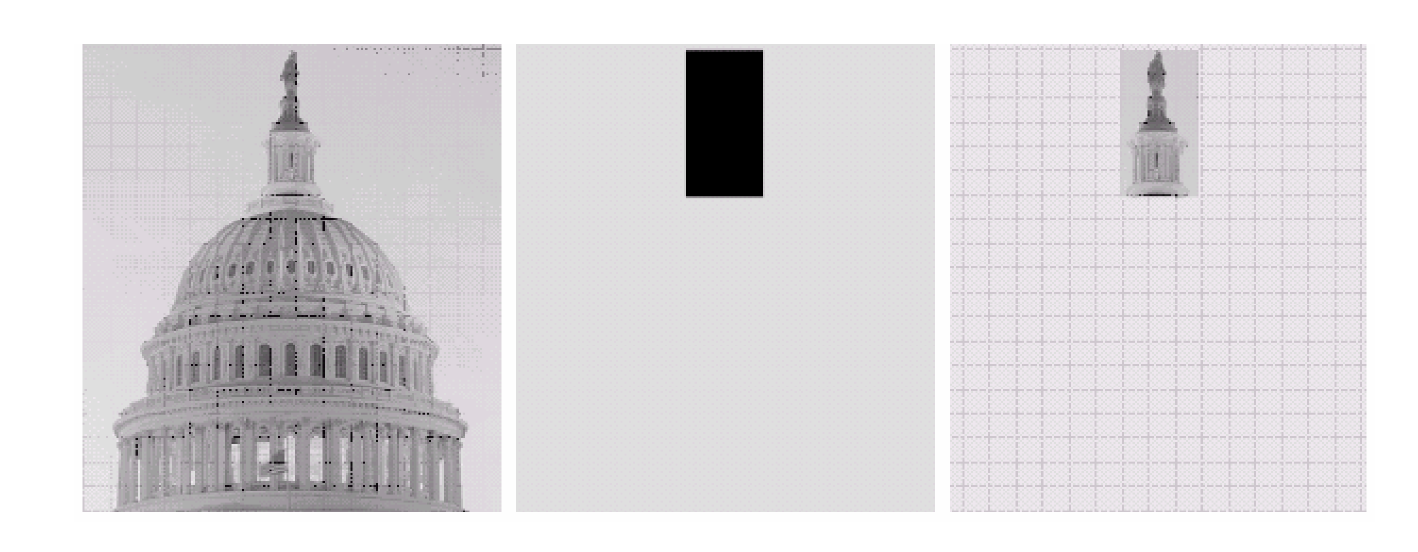
- 主要应用举例
  - ✓ 合并子图像





代数运算——3、或

- ✓ 模板运算：提取感兴趣的子图像



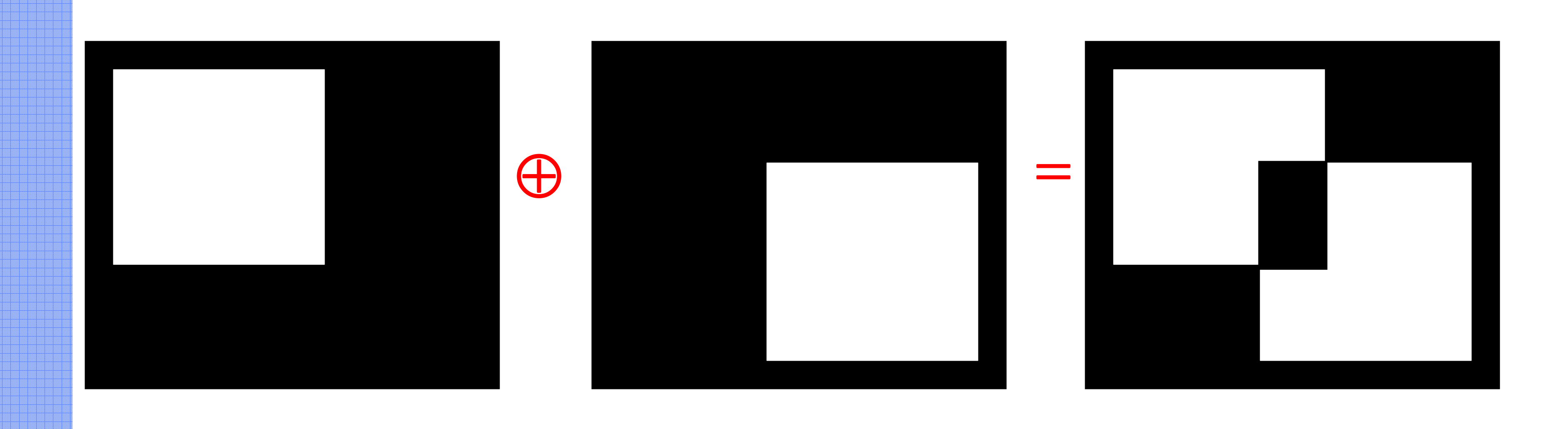
代数运算——4、异或



- 异或运算的定义

  $$g(x, y) = f(x, y) \oplus h(x, y)$$

- 主要应用举例
  - ✓ 获得相交子图像



直方图运算



- 直方图定义
- 直方图均衡化

直方图定义



- 图像直方图的定义（1）

一个灰度级在范围[0，L-1]的数字图像的直方图是一个离散函数

$$h(r_k) = n_k$$

n_k是图像中灰度级为r_k的像素个数

r_k 是第k个灰度级，$k = 0, 1, 2, \ldots, L-1$

由于r_k的增量是1，直方图可表示为：

$$p(k) = n_k$$

即，图像中不同灰度级像素出现的次数

直方图定义



- 图像直方图的定义（2）

  一个灰度级在范围[0，L-1]的数字图像的直方图是一个离散函数

$$p(r_k) = n_k/n$$

n 是图像的像素总数

n_k是图像中灰度级为r_k的像素个数

r_k 是第k个灰度级，$k = 0, 1, 2, \ldots, L-1$

直方图定义



- 两种图像直方图定义的比较

$$h(r_k) = n_k \qquad \text{定义(1)}$$

$$p(r_k) = n_k/n \qquad \text{定义(2)}$$

其中，定义（2）

✓ 使函数值正则化到[0,1]区间，成为实数函数

✓ 函数值的范围与象素的总数无关

✓ 给出灰度级r_k在图像中出现的概率密度统计



直方图定义

- 图像直方图的定义举例

$p(r_k)$

r_k



像素数目$p(r_k)$

Dark image

灰度级r_k

Bright image

Low-contrast image

High-contrast image

直方图均衡化达到的效果

直方图均衡化



- 直方图应用举例——直方图均衡化
  - ✓ 希望一幅图像的像素占有全部可能的灰度级且分布均匀，能够具有高对比度
  - ✓ 使用的方法是灰度级变换：$s = T(r)$
  - ✓ 基本思想是把原始图的直方图变换为均匀分布的形式，这样就增加了像素灰度值的动态范围，从而达到增强图像整体对比度的效果

直方图均衡化



$s=T(r) \qquad 0 \leqslant r \leqslant 1$

T(r)满足下列两个条件：

（1）T(r)在区间$0 \leqslant r \leqslant 1$中为单值且单调递增

（2）当$0 \leqslant r \leqslant 1$时，$0 \leqslant T(r) \leqslant 1$

条件（1）保证原图各灰度级在变换后仍保持从黑到白（或从白到黑）的排列次序

条件（2）保证变换前后灰度值动态范围的一致性

直方图均衡化



$P_r(r)$是r的概率密度函数，$P_s(s)$是s的概率密度函数，$P_r(r)$和$T(r)$已知，且$T^{-1}(s)$满足上述条件(1)，所以

$$P_s(s) = P_r(r)\left|\frac{dr}{ds}\right|$$

直方图均衡化



已知一种重要的变换函数：$s = T(r) = \int_0^r p_r(w)dw$

关于上限的定积分的导数就是该上限的积分值（莱布尼茨准则）

$$\frac{ds}{dr} = \frac{dT(r)}{dr} = \frac{d}{dr}\left[\int_0^r p_r(w)dw\right] = p_r(r)$$

$$p_s(s) = p_r(r)\left|\frac{dr}{ds}\right| = p_r(r)\left|\frac{1}{p_r(r)}\right| = 1$$

直方图均衡化



对于离散值：$p_r(r_k) = \frac{n_k}{n}$

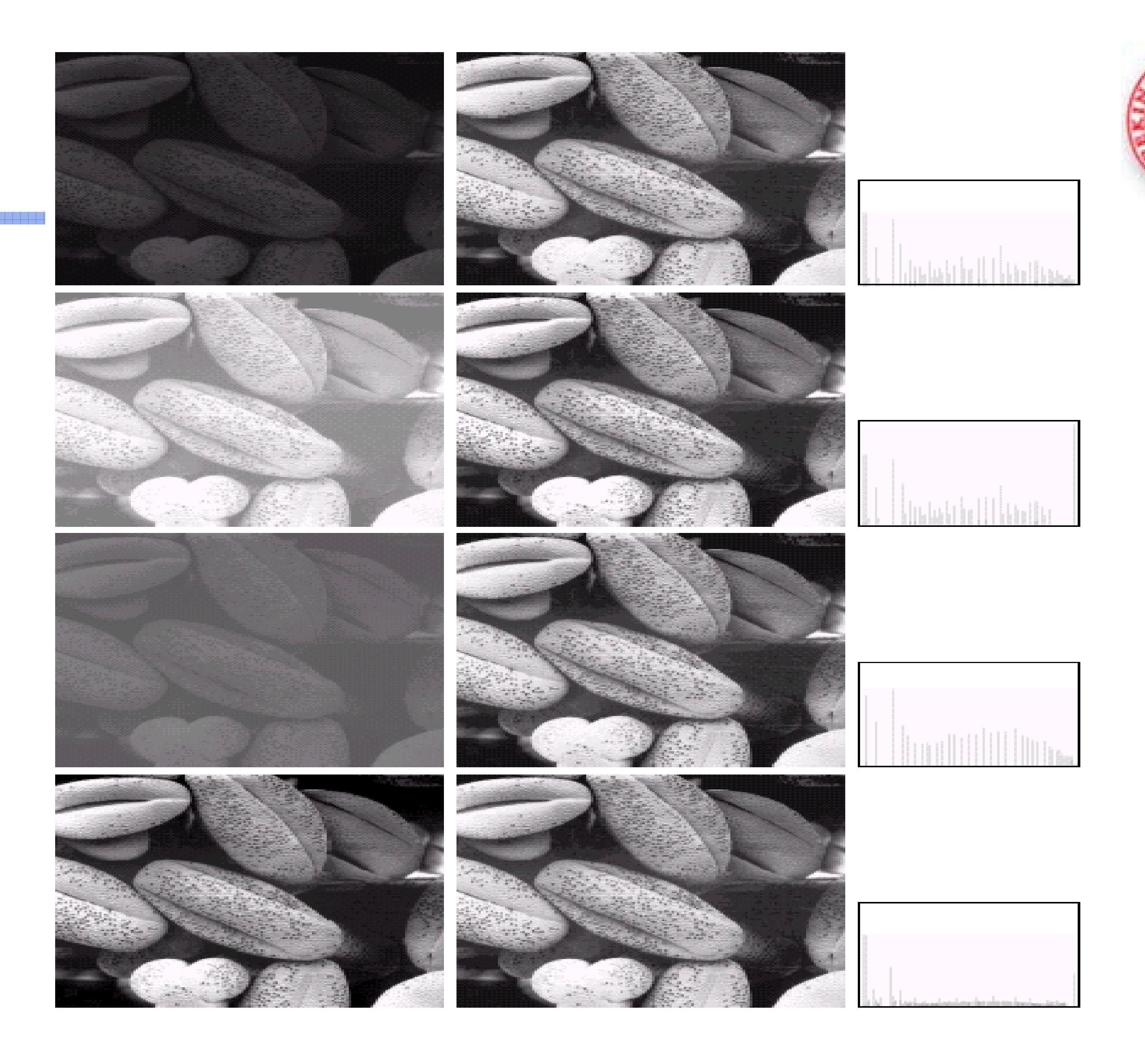
已知变换函数的离散形式为：

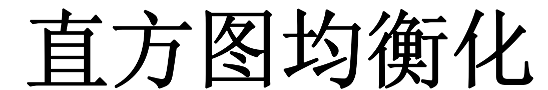
$$s_k = T(r_k) = \sum_{j=0}^{k} p_r(r_j) = \sum_{j=0}^{k} \frac{n_j}{n}$$

k＝0, 1, 2, …, L-1

s_k称作直方图均衡化

将输入图像中灰度级为r_k（横坐标）的像素映射到输出图像中灰度级为s_k（横坐标）的对应像素得到



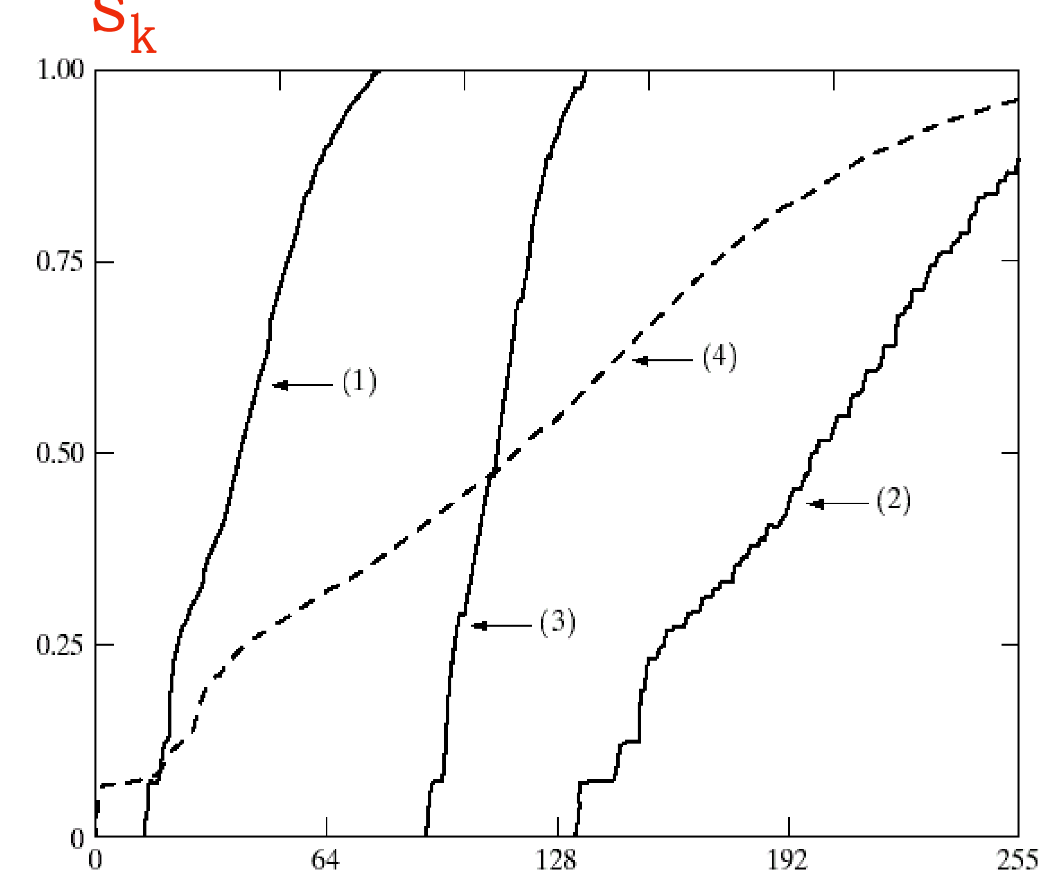


直方图均衡化



$$s_k = T(r_k) = \sum_{j=0}^{k} p_r(r_j) = \sum_{j=0}^{k} \frac{n_j}{n}$$

FIGURE 3.18 Transformation functions (1) through (4) were obtained from the histograms of the images in Fig.3.17(a), using Eq. (3.3-8).



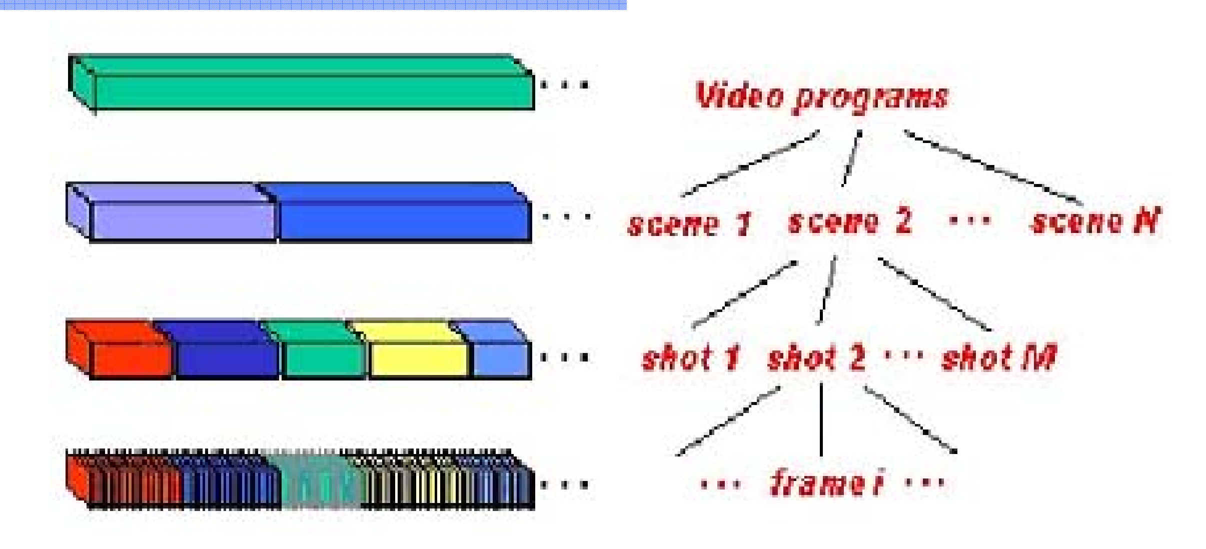
应用——镜头边界的检测



- 为什么要补充该部分知识?
  - ✓ 涉及两幅图像相减
  - ✓ 涉及直方图知识
  - ✓ 涉及两幅图像直方图的相减
- 什么是镜头?
- 镜头检测方法1——连续帧相减
- 镜头检测方法2——连续帧的直方图相减
- 镜头检测方法3——时空切片分析

什么是镜头？





- 从视频的制造产生来看，视频由一个个镜头(shot)所组成
- 一个镜头是指一系列连续记录的图像帧，用于表示一个时间段或相同地点连续的动作
- 镜头由摄像机一次摄像的开始和结束所决定

什么是镜头？



- 检测到镜头及由镜头产生的关键帧，可以：

  - ✓ 提供基于关键帧的视频浏览
  - ✓ 提供基于内容的视频检索和查询
  - ✓ 计算机自动分析和总结，节省人力和时间
  - ✓ 使海量视频数据的管理和索引成为可能

镜头检测方法1——连续帧相减



- **算法原理：**计算相邻两帧像素变化的数目。当超过设定的阈值时，即找到镜头的边界
- **缺点：**对摄像机运动敏感，如放缩、平移
- **解决办法：**通过滤波器的使用来降低。在比较一帧的每个像素前，用它的邻近区域的平均值来代替，这也过滤了输入图像的一些噪声



√



×



√



√

镜头检测方法2——直方图相减



- **算法原理：**统计相邻两帧中所有像素在不同灰度（颜色）上的分布差异，当差异的累加值超过阈值T时，即检测到镜头边界（后面有算法的详细说明）
- **优点：**对对象运动不敏感，因为直方图忽略了帧内的空间变化
- **缺点：**可能两个图像有类似的直方图但却是完全不同的内容。然而，这种事件的概率是足够低

镜头检测方法2——直方图相减



（1）首先计算相邻两帧颜色分布差值的均值和方差

$$\mu = \frac{1}{M}\sum_{i=1}^{M} d_i \qquad (1)$$

$$\sigma = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(d_i - \mu)^2} \qquad (2)$$

d_i 表示第 i 帧和第 i+1 帧的区别，它通过下式计算

$$d_i = \sum_{n=1}^{N}\left|H_i(n) - H_{i+1}(n)\right| \qquad (3)$$

$H_i(n)$ 表示第 i 帧直方图的值：

$$H_i(n) = \frac{C_n}{C} \qquad (4)$$

C_n 是第 i 帧中 P(x, y) =n 的像素数目，C 是第 i 帧的总像素数。n 是 N 种可能颜色的一种。N=64，采用每个像素点的 RGB 三种色的头两位组合而成。

$$P(x,y) = ((R(x,y) >> 6) << 4) + ((G(x,y) >> 6) << 2) + ((B(x,y) >> 6) \qquad (5)$$

镜头检测方法2——直方图相减



（2）、根据 μ 和 σ，计算 threshold T

$$T_b = \mu + 3.8 \times \sigma \quad (6)$$

$$T_s = \mu \times 1\ 1 \quad (7)$$

当 $d_i > T_b$, cut 被检测；

当 $T_s < d_i < T_b$, 可能出现渐进转换，当帧间积累的变化值超过 T_b, 则检测到镜头边界；

当 $T_s > d_i$, 没有出现镜头边界的可能。

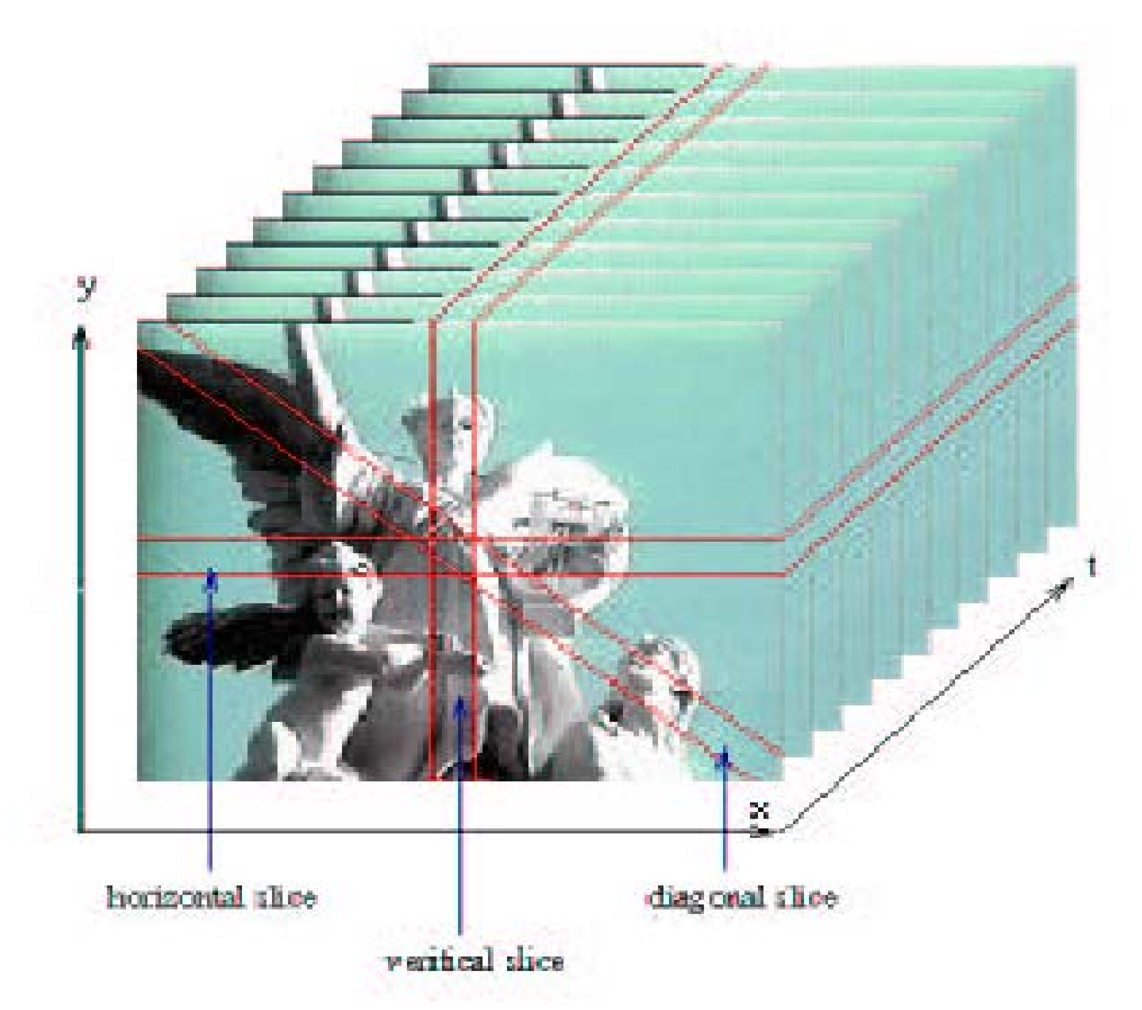
参考文献：

H.J.Zhang,A.Kankanhalli,and S.W.Smoliar, “Automatic Partitioning of Full-Motion Video”. ACM Multimedia System, Apr.1993.

镜头检测方法3——时空切片分析



- 什么是时空切片spatio-temporal slices?



镜头检测方法3——时空切片分析



- 颜色和纹理的不连续，表明镜头边界的出现

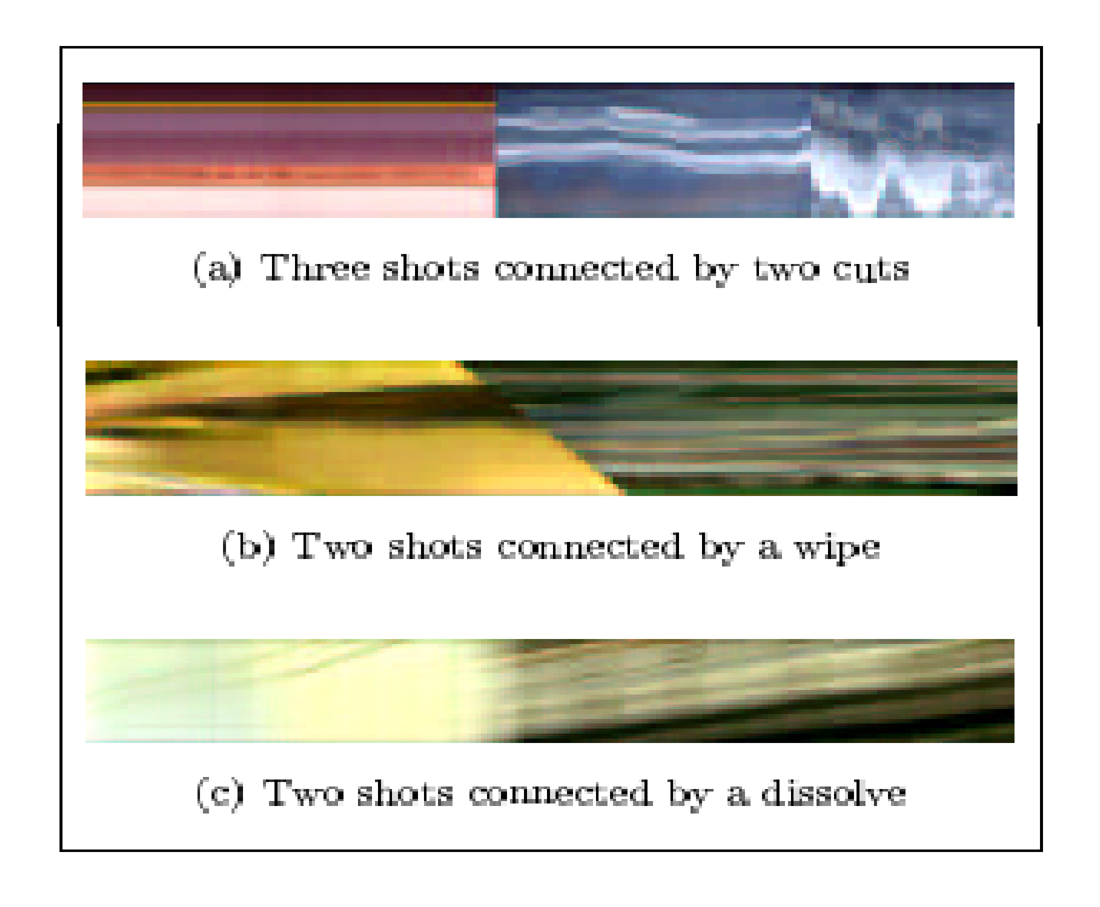


Figure 3.4: Samples of spatio-temporal slices.

镜头检测方法3——时空切片分析



- 纹理的方向表明摄像机和对象的不同运动

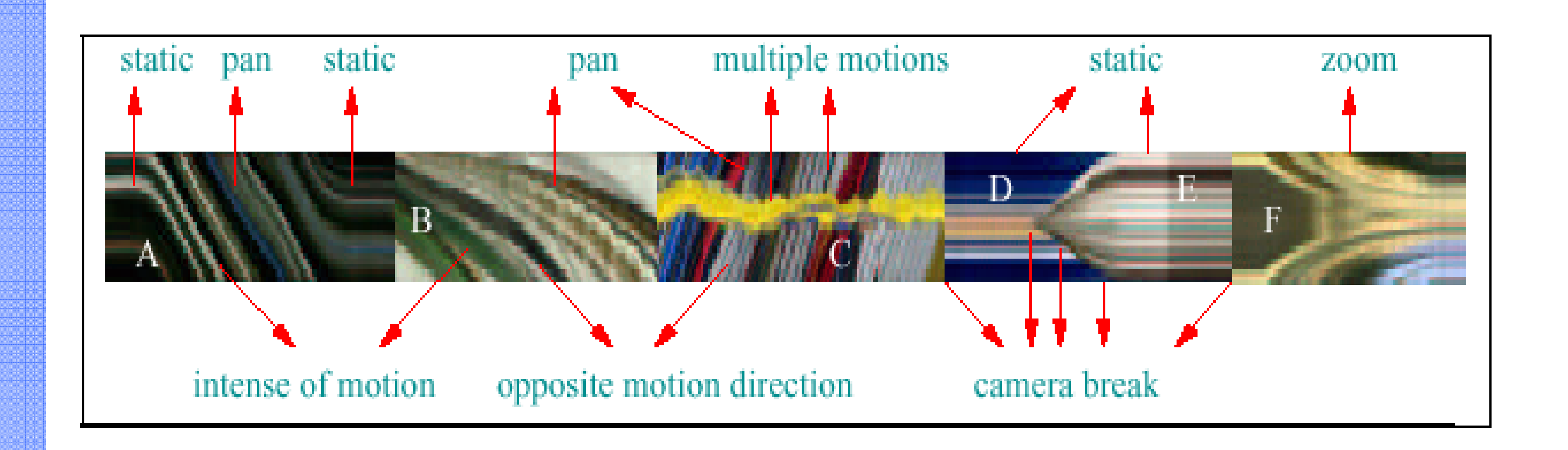


Figure 2.3: Patterns in a spatio-temporal slice.

镜头检测方法3——时空切片分析



- 参考文献：

C.W.Ngo,T.C.Pong,and R.T.Chin. “Video Partitioning by Temporal Slice Coherency”. IEEE Transactions on Circuits and Systems for Video Technology, Aug.2001.

空间域图像增强



- 基础知识
  - ✓ 基本概念
  - ✓ 点运算
  - ✓ 代数运算
  - ✓ 直方图运算
  - ✓ 应用－镜头边界的检测
- 空间滤波器
  - ✓ 平滑空间滤波器
  - ✓ 锐化空间滤波器

空间滤波器



- 空间滤波和空间滤波器的定义

使用空间模板进行的图像处理，被称为空间滤波。模板本身被称为空间滤波器

空间滤波和空间滤波器的定义



- 在 $M \times N$ 的图像f上，使用 $m \times n$ 的滤波器：

$$g(x,y)=\sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t) f(x+s,y+t)$$

其中，m=2a+1, n=2b+1,

w(s, t)是滤波器系数，f(x, y)是图像值

- 空间滤波的简化形式：

$$R=w_1 z_1+w_2 z_2+\ldots+w_{mn} z_{mn}=\sum_{i=1}^{mn} w_i z_i$$

其中，w是滤波器系数，z是与该系数对应的图像灰度值，mn为滤波器中包含的像素点总数



- 平滑空间滤波器的作用
  - 模糊处理:去除图像中一些不重要的细节
  - 减小噪声
- 平滑空间滤波器的分类
  - 线性滤波器：均值滤波器
  - 非线性滤波器
    - 最大值滤波器
    - 中值滤波器
    - 最小值滤波器

线性滤波器



- 包含在滤波器邻域内像素的平均值，也称为均值滤波器
- 作用
  - ✓ 减小图像灰度的“尖锐”变化，减小噪声
  - ✓ 由于图像边缘是由图像灰度尖锐变化引起的，所以也存在边缘模糊的问题

线性滤波器



a

$$\frac{1}{9}\times\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$$

b

$$\frac{1}{16}\times\begin{bmatrix}1&2&1\\2&4&2\\1&2&1\end{bmatrix}$$

图a是标准的像素平均值

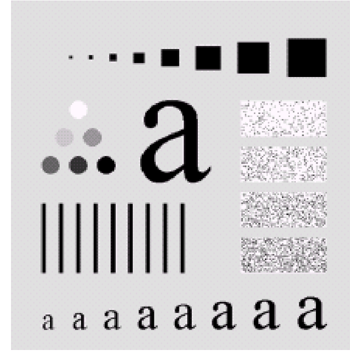
图b是像素的加权平均，表明一些像素更为重要

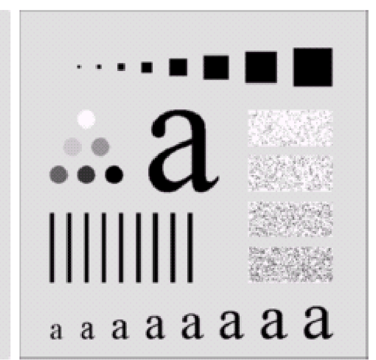
$$g(x,y)=\frac{\sum_{s=-a}^{a}\sum_{t=-b}^{b}w(s,t)f(x+s,y+t)}{\sum_{s=-a}^{a}\sum_{t=-b}^{b}w(s,t)}$$

线性滤波器——例1



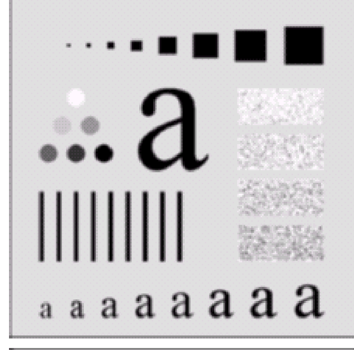


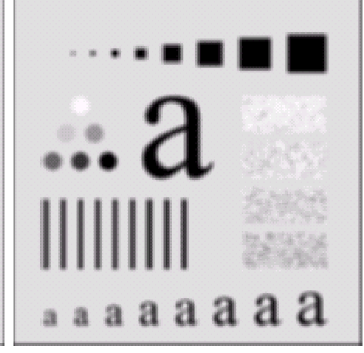




3 x 3

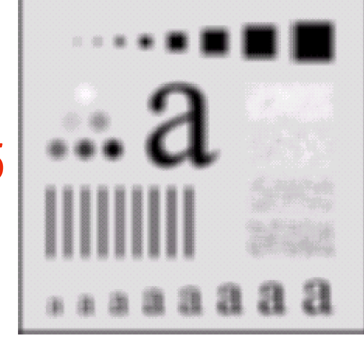
5 x 5

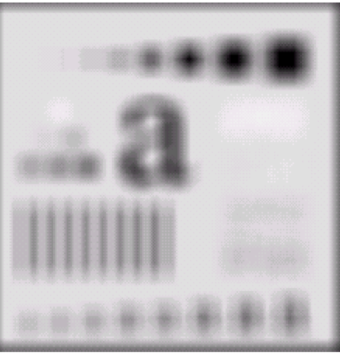




9 x 9

15 x 15





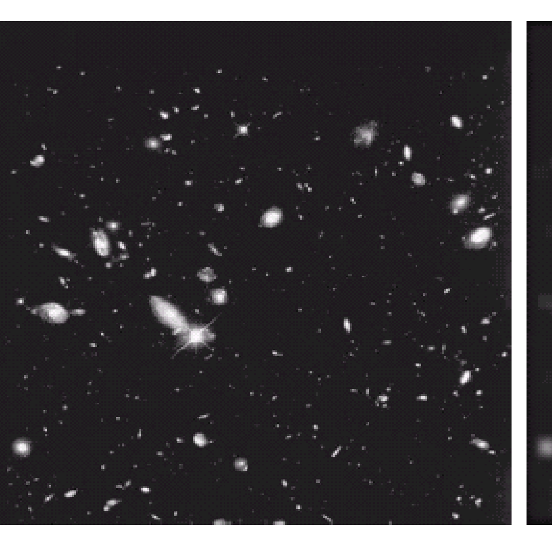
35 x 35

线性滤波器——例2

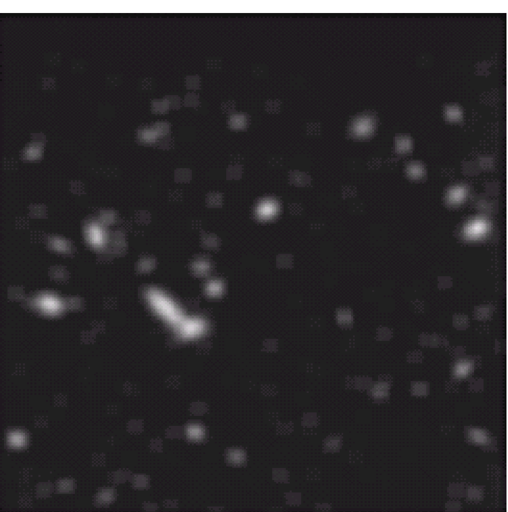


提取感兴趣物体而模糊图像

原图



15 x 15



阈值＝25% x b图像的最高亮度



a b c

FIGURE 3.36 (a) Image from the Hubble Space Telescope. (b) Image processed by a 15 × 15 averaging mask. (c) Result of thresholding (b). (Original image courtesy of NASA.)

统计排序滤波器



- 什么是统计排序滤波器？
  - ✓ 是一种非线性滤波器
  - ✓ 基于滤波器所在图像区域中像素的排序，由排序结果决定的值代替中心像素的值
- 分类
  - ✓ 中值滤波器： 用像素领域内的中间值代替该像素
  - ✓ 最大值滤波器：用像素领域内的最大值代替该像素
  - ✓ 最小值滤波器：用像素领域内的最小值代替该像素

统计排序滤波器



- 中值滤波器
  - ✓ 主要用途：去除噪声
  - ✓ 计算公式：$R = \text{mid}\ \{z_k \mid k = 1, 2, \ldots, n\}$
- 最大值滤波器
  - ✓ 主要用途：寻找最亮点
  - ✓ 计算公式：$R = \max\ \{z_k \mid k = 1, 2, \ldots, n\}$
- 最小值滤波器
  - ✓ 主要用途：寻找最暗点
  - ✓ 计算公式：$R = \min\ \{z_k \mid k = 1, 2, \ldots, n\}$

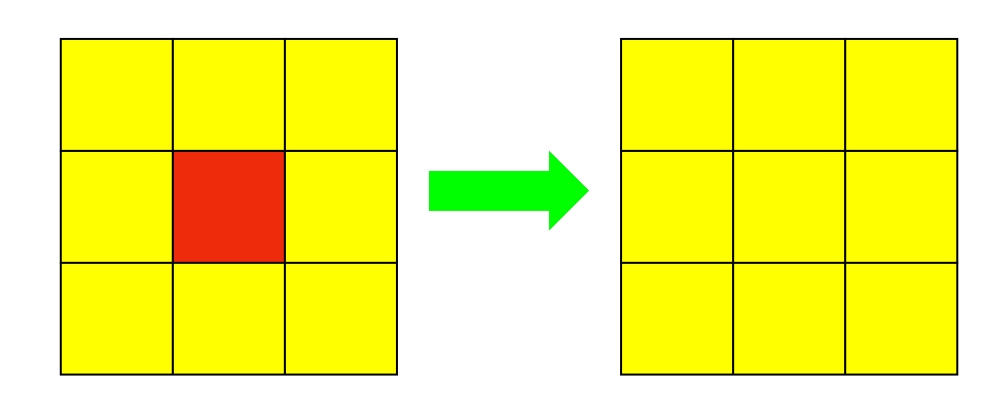
中值滤波器



- 中值滤波的原理
  - ✓ 用模板区域内像素的中间值，作为结果值

$$R = \text{mid}\ \{z_k \mid k = 1, 2, \ldots, n\}$$

  - ✓ 强迫突出的亮点（暗点）更象它周围的值，以消除孤立的亮点（暗点）



中值滤波器



- 中值滤波算法的实现

✓将模板区域内的像素排序，求出中间值

例如：3x3的模板，第5大的是中值，

5x5的模板，第13大的是中值，

7x7的模板，第25大的是中值，

9x9的模板，第41大的是中值。

✓对于同值像素，连续排列。

如（10, 15, 20, 20, 20, 20, 20, 25, 100）

中值滤波器

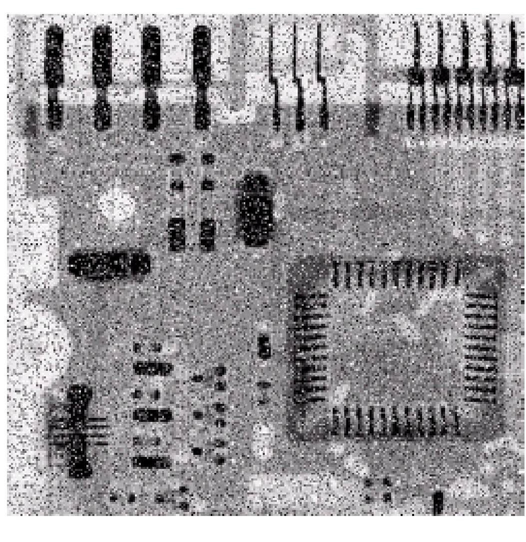


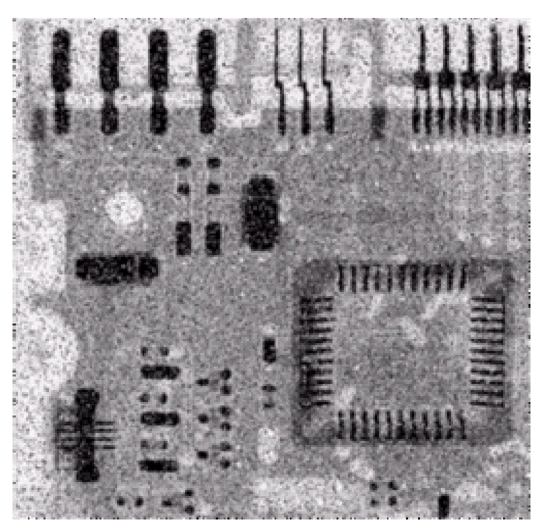
- 中值滤波算法的特点
  - ✓在去除噪音的同时，可以比较好地保留边的锐度和图像的细节（优于均值滤波器）
  - ✓能够有效去除脉冲噪声：以黑白点叠加在图像上

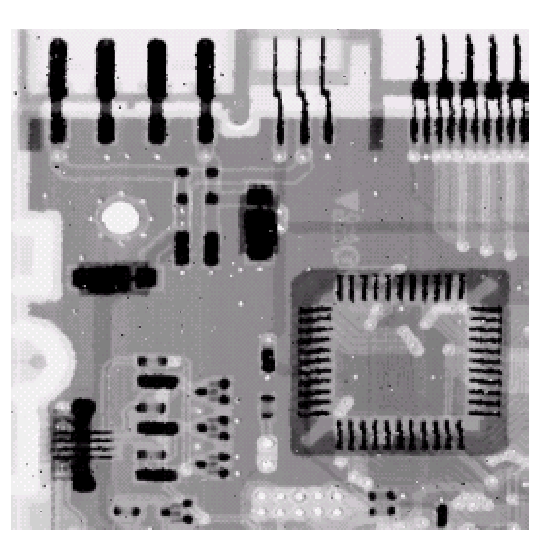
中值滤波器



原图 3x3均值滤波 3x3中值滤波





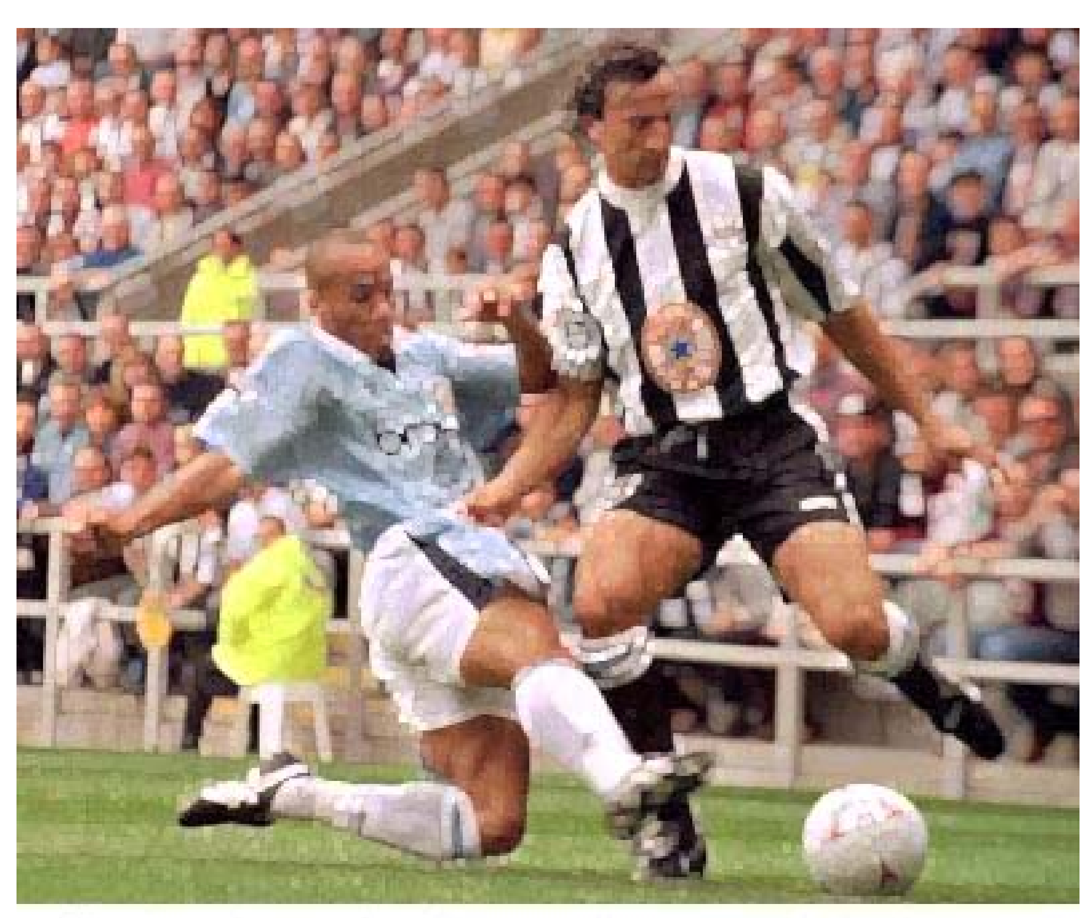


a b c

FIGURE 3.37 (a) X-ray image of circuit board corrupted by salt-and-pepper noise. (b) Noise reduction with a 3×3 averaging mask. (c) Noise reduction with a 3×3 median filter. (Original image courtesy of Mr. Joseph E. Pascente, Lixi, Inc.)

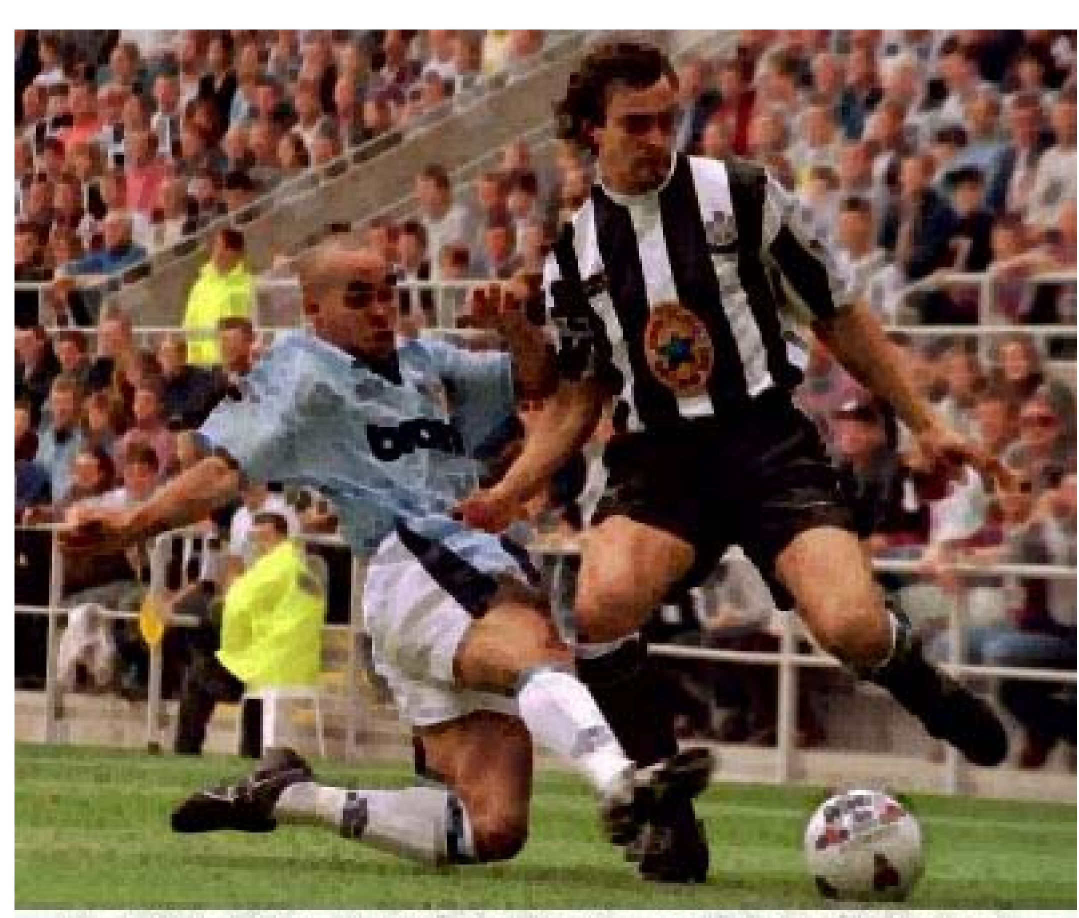
最大值滤波器





最小值滤波器





Man City's Richard Edghill tackles Newcastle's David Ginola at St James Park. City lost the match 3-1 Picture: REUTERS

锐化滤波器



- 锐化滤波器的主要用途
  - ✓ 突出图像中的细节，增强被模糊了的细节
  - ✓ 印刷中的细微层次强调。弥补扫描对图像的钝化
  - ✓ 超声探测成像，分辨率低，边缘模糊，通过锐化来改善
  - ✓ 图像识别中，分割前的边缘提取
  - ✓ 锐化处理恢复过度钝化、暴光不足的图像
  - ✓ 尖端武器的目标识别、定位

锐化滤波器



- 微分滤波器的原理
  - 均值产生钝化的效果，而均值与积分相似，由此而联想到，微分能不能产生相反的效果，即锐化的效果？结论是肯定的。
  - 在图像处理中应用微分最常用的方法是计算梯度。函数f(x, y)在(x, y)处的梯度为一个向量：

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

锐化滤波器



- 锐化滤波器的分类
  - ✓ 二阶微分滤波器－拉普拉斯算子
  - ✓ 一阶微分滤波器－梯度算子

拉普拉斯算子



- 图像函数的拉普拉斯变换定义为

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

$$\frac{\partial^2 f}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y)$$

$$\frac{\partial^2 f}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y)$$

$$\nabla^2 f = [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1)] - 4f(x, y)$$

拉普拉斯算子



0	1	0
1	−4	1
0	1	0

1	1	1
1	−8	1
1	1	1

0	−1	0
−1	4	−1
0	−1	0

−1	−1	−1
−1	8	−1
−1	−1	−1

a	b
c	d

FIGURE 3.39 (a) Filter mask used to implement the digital Laplacian, as defined in Eq. (3.7-4). (b) Mask used to implement an extension of this equation that includes the diagonal neighbors. (c) and (d) Two other implementations of the Laplacian.

拉普拉斯算子



- 拉普拉斯变换对图像增强的基本方法

$$g(x,y)=\begin{cases} f(x,y)-\nabla^2 f(x,y) & (1) \\ f(x,y)+\nabla^2 f(x,y) & (2) \end{cases}$$

(1)用于拉普拉斯模板中心系数为负

(1)用于拉普拉斯模板中心系数为正

拉普拉斯算子一例

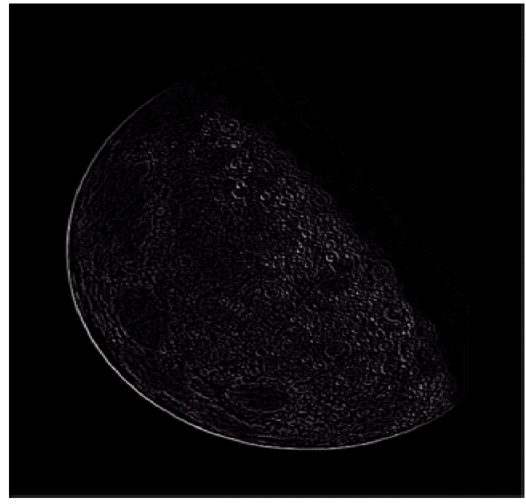


原图:月球北极　　拉普拉斯滤波后的图像

a	b
c	d

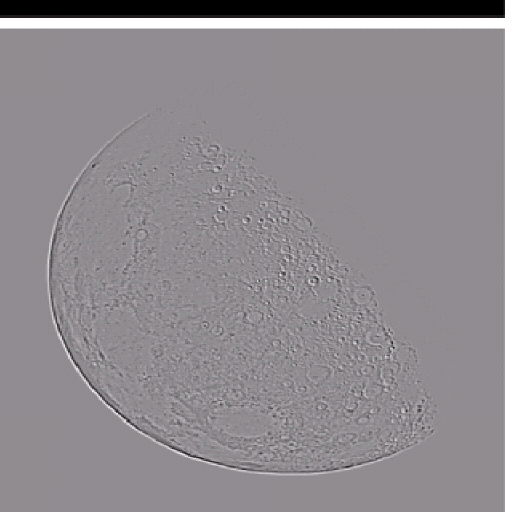
FIGURE 3.40
(a) Image of the North Pole of the moon.
(b) Laplacian-filtered image.
(c) Laplacian image scaled for display purposes.
(d) Image enhanced by using Eq. (3.7-5).
(Original image courtesy of NASA.)





3×3，中心点为－8的掩膜

标定的图像





原始图像＋拉普拉斯的结果

拉普拉斯算子



$$g(x,y)=f(x,y)-\nabla^2 f(x,y)$$

$$=f(x,y)-\{[f(x+1,y)+f(x-1,y+f(x,y+1)+f(x,y-1))]-4f(x,y)\}$$

$$=5f(x,y)-[f(x+1,y)+f(x-1,y)+f(x,y+1)+f(x,y-1)]$$

$$g(x,y)=f(x,y)+\nabla^2 f(x,y)$$

$$=f(x,y)+\{[f(x+1,y)+f(x-1,y+f(x,y+1)+f(x,y-1))]-4f(x,y)\}$$

$$=[f(x+1,y)+f(x-1,y)+f(x,y+1)+f(x,y-1)]-3f(x,y)$$

拉普拉斯算子一例



a

0	−1	0
−1	5	−1
0	−1	0

b

−1	−1	−1
−1	9	−1
−1	−1	−1

原图

a 滤波

b 滤波

比 a 更锐化

a b c
d e

FIGURE 3.41 (a) Composite Laplacian mask. (b) A second composite mask. (c) Scanning electron microscope image. (d) and (e) Results of filtering with the masks in (a) and (b), respectively. Note how much sharper (e) is than (d). (Original image courtesy of Mr. Michael Shaffer, Department of Geological Sciences, University of Oregon, Eugene.)

梯度算子



- 梯度通过一个二维列向量来定义

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

- 向量的模值

$$\nabla f = mag(\nabla f)$$

$$= \left[G_x^2 + G_y^2 \right]^{\frac{1}{2}}$$

$$= \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right]^{\frac{1}{2}}$$

梯度算子



- 考虑一个3x3的图像区域，z代表灰度级，上式在点z_5的∇f值可用数字方式近似。

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

$G_x = \frac{\partial f}{\partial x}$ 用（$z_6 - z_5$）近似

$G_y = \frac{\partial f}{\partial y}$ 用（$z_8 - z_5$）近似，

组合为：

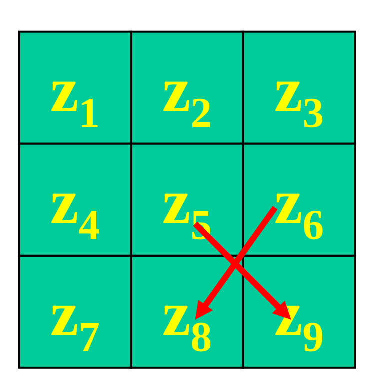
$$\nabla f \approx \left[(z_6 - z_5)^2 + (z_8 - z_5)^2 \right]^{\frac{1}{2}}$$

梯度算子



- 向量模值的近似计算

用绝对值替换平方和平方根有：



$$\nabla f = \left[\left| G_x^2 \right| + \left| G_y^2 \right| \right]^{\frac{1}{2}}$$

$$\approx \left| G_x \right| + \left| G_y \right|$$

- 微分过滤器的原理

$$\nabla f \approx \left| z_6 - z_5 \right| + \left| z_8 - z_5 \right|$$

另外一种计算方法是使用交叉差：

$$\nabla f \approx \left[(z_9 - z_5)^2 + (z_8 - z_6)^2 \right]^{1/2}$$

$$\approx \left| z_9 - z_5 \right| + \left| z_8 - z_6 \right|$$

梯度算子



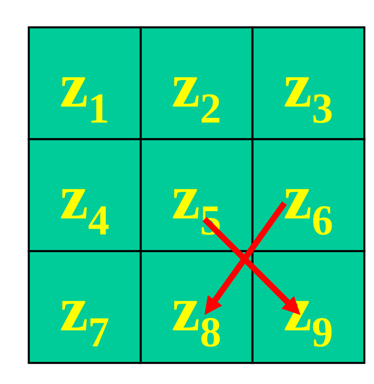
- 微分滤波器模板系数设计
  - ✓Roberts交叉梯度算子
  - ✓Prewitt梯度算子
  - ✓Sobel梯度算子

微分滤波器模板系数设计



- Roberts交叉梯度算子

$$\nabla f \approx |z_9 - z_5| + |z_8 - z_6|$$



✓梯度计算由两个模板组成，第一个求得梯度的第一项，第二个求得梯度的第二项，然后求和，得到梯度。

✓两个模板称为Roberts交叉梯度算子

-1	0
0	1

0	-1
1	0

微分滤波器模板系数设计



- Prewitt梯度算子——3x3的梯度模板

$$\nabla f \approx |(z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)| + |(z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)|$$

-1	-1	-1
0	0	0
1	1	1

-1	0	1
-1	0	1
-1	0	1

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

微分滤波器模板系数设计



- Sobel梯度算子——3x3的梯度模板

$$\nabla f \approx |(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)| + |(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)|$$

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

数字图像处理（2）



任何问题？

研究生课程



数字图像处理

Digital Image Processing

彭宇新

北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

彩色图像处理



- 彩色基础知识
- 彩色空间
- 伪彩色处理
- 全彩色图像处理
- 彩色变换
- 彩色图像平滑和尖锐化

彩色图像基础



- 为什么要研究彩色图像处理？
  - ✓ 符合人类视觉特点
    - ➢ 人类可以辨别几千种颜色色调和亮度
    - ➢ 只能辨别几十种灰度层次
  - ✓ 有用的描绘子
    - ➢ 简化目标物的区分
    - ➢ 目标识别：根据目标的颜色特征

彩色图像基础

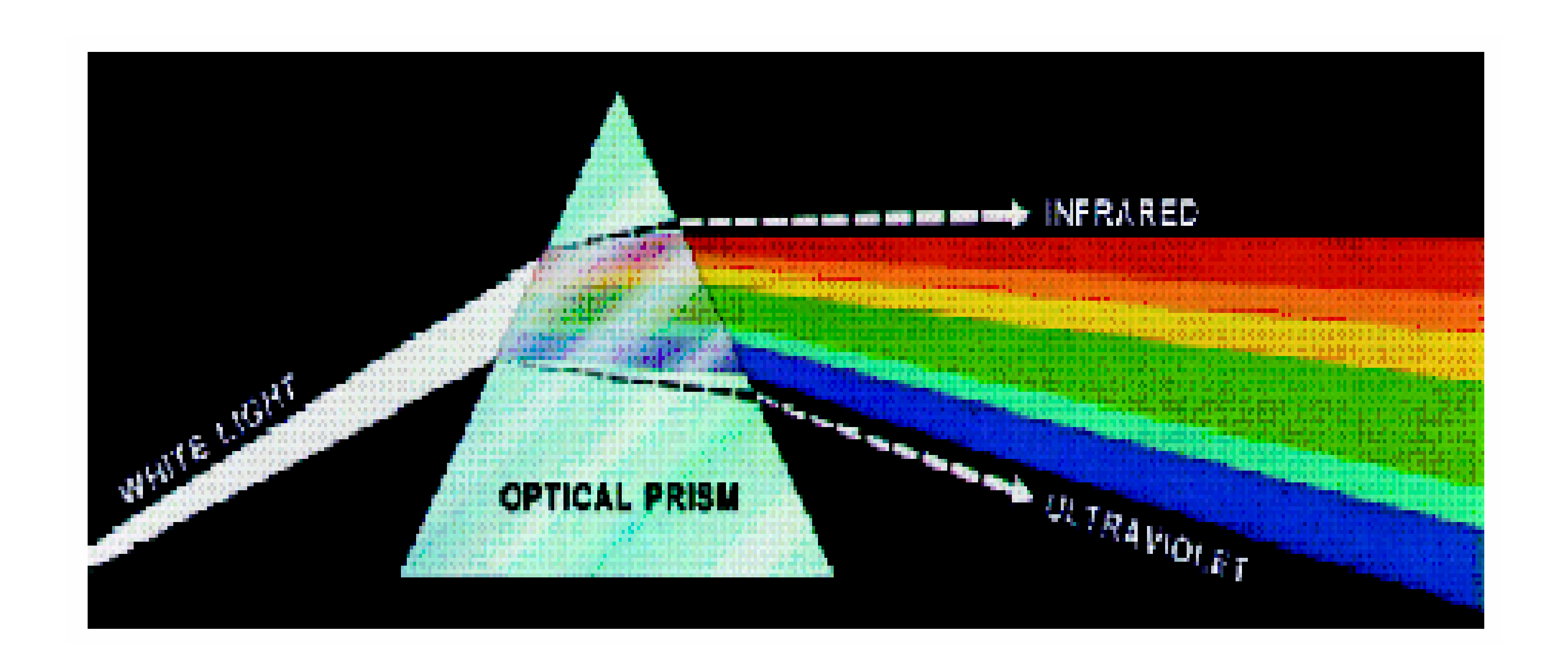


- 彩色图像处理可分为：
  - ✓ 全彩色处理
    - 数码相机
    - 数码摄像机
    - 彩色扫描仪
  - ✓ 伪彩色处理
    - 对不同的灰度或灰度范围赋予不同的颜色

彩色图像基础



当一束白光通过一个玻璃棱镜时，出现的光束不是白光，而是由一端为紫色到另一端为红色的连续彩色谱组成



彩色图像基础



- 光特性是颜色科学的核心
- 描述彩色光的3个基本量：
  - ✓ 辐射率：从光源流出能量的总量，用瓦特(W)度量
  - ✓ 光强：观察者从光源接收的能量总和
  - ✓ 亮度：主观描绘子

彩色图像基础



- 三原色
  - ✓红色（Red）、绿色（Green）、蓝色（Blue）
- 原色相加可产生二次色
  - ✓深红色：红＋蓝
  - ✓青　色：绿＋蓝
  - ✓黄　色：红＋绿

彩色图像处理



- 彩色基础知识
- 彩色空间
- 伪彩色处理
- 全彩色图像处理
- 彩色变换
- 彩色图像平滑和尖锐化

彩色空间（也称彩色模型或彩色系统）



- RGB
- CMY和CMYK
- HSI
- YIQ
- YUV
- YCbCr

彩色空间（也称彩色模型或彩色系统）



- RGB
  - CCD技术直接感知R,G,B三个分量
  - 是图像成像、显示、打印等设备的基础

CMY和CMYK彩色空间



- CMY(青、深红、黄)、CMYK（青、深红、黄、黑）
  - ✓ 运用在大多数在纸上沉积彩色颜料的设备，如彩色打印机和复印机
  - ✓ CMYK
    - ➢ 打印中的主要颜色是黑色
    - ➢ 等量的CMY原色产生黑色，但不纯
    - ➢ 在CMY基础上，加入黑色，形成CMYK彩色空间

彩色空间（也称彩色模型或彩色系统）



- HSI(色调、饱和度、亮度）
  - ✓ 两个特点：
    - ➢ I分量与图像的彩色信息无关
    - ➢ H和S分量与人感受颜色的方式是紧密相连的
  - ✓ 将亮度(I)与色调（H）和饱和度（S）分开
  - ✓ 避免颜色受到光照明暗(I)等条件的干扰
  - ✓ 仅仅分析反映色彩本质的色调和饱和度
  - ✓ 广泛用于计算机视觉、图像检索和视频检索

彩色空间（也称彩色模型或彩色系统）



- YIQ
  - ✓ Y指亮度(Brightness),即灰度值
  - ✓ I和Q指色调，描述色彩及饱和度
  - ✓ 用于彩色电视广播，被北美的电视系统所采用（属于NTSC系统）
  - ✓ Y分量可提供黑白电视机的所有影像信息

彩色空间（也称彩色模型或彩色系统）



- YUV
  - ✓ Y指亮度,与YIQ的Y相同
  - ✓ U和V也指色调，不同于YIQ的I和Q
  - ✓ 用于彩色电视广播，被欧洲的电视系统所采用（属于PAL系统）
  - ✓ Y分量也可提供黑白电视机的所有影像信息

彩色空间（也称彩色模型或彩色系统）



- YCbCr
  - Y指亮度,与YIQ和YUV的Y相同
  - Cb和Cr由U和V调整得到
  - JPEG采用的彩色空间

彩色空间转换



1. RGB ⟷ CMY
2. RGB ⟷ HSI
3. RGB ⟷ YIQ
4. RGB ⟷ YUV
5. RGB ⟷ YCbCr

1、RGB ⟷ CMY



$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

RGB和CMY值都归一化到[0, 1]

2、RGB ⟶ HSI



$$H = \begin{cases} \theta & B \le G \\ 360 - \theta & B > G \end{cases}$$

$$\theta = \arccos\left\{ \frac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\left[(R-G)^2+(R-G)(G-B)\right]^{\frac{1}{2}}} \right\}$$

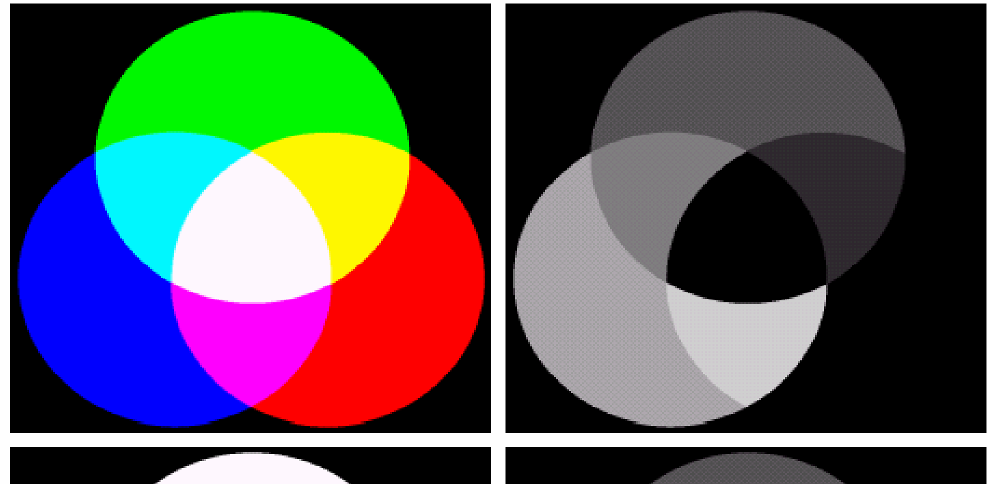
$$S = 1 - \frac{3}{(R+G+B)}\left[\min(R,G,B)\right]$$

$$I = \frac{1}{3}(R + G + B)$$

RGB图像和与之对应的HSI图像分量

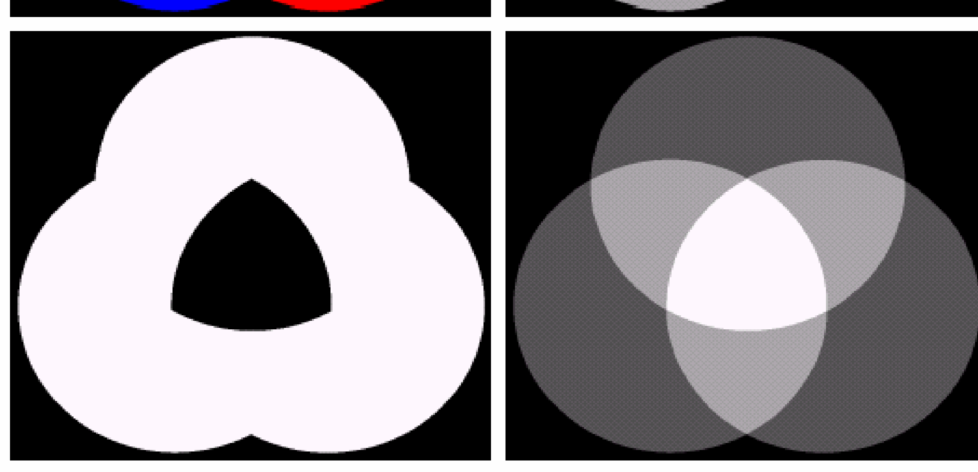


RGB图像



色调

饱和度



强度

a	b
c	d

FIGURE 6.16 (a) RGB image and the components of its corresponding HSI image: (b) hue, (c) saturation, and (d) intensity.

2、HSI ⟶ RGB



$$0^0 \le H < 120^0$$

$$R = I\left[1 + \frac{S\cos H}{\cos\left(60^0 - H\right)}\right] \qquad G = 1 - (R + B) \qquad B = I(1 - s)$$

$$120^0 \le H < 240^0 \qquad H = H - 120^0$$

$$G = I\left[1 + \frac{S\cos H}{\cos\left(60^0 - H\right)}\right] \qquad B = 1 - (R + G) \qquad R = I(1 - s)$$

$$240^0 \le H < 360^0 \qquad H = H - 240^0$$

$$B = I\left[1 + \frac{S\cos H}{\cos\left(60^0 - H\right)}\right] \qquad R = 1 - (G + B) \qquad G = I(1 - s)$$

3、RGB ⟷ YIQ



$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0.956 & 0.621 \\ 1 & -0.272 & -0.647 \\ 1 & -1.106 & 1.703 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$$

4、RGB ⟷ YUV



$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.148 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.140 \\ 1 & -0.395 & -0.581 \\ 1 & 2.032 & 0 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix}$$

5、RGB ⟷ YCbCr



$$\begin{bmatrix} Y \\ Cb \\ Cr \\ 1 \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 & 0 \\ -0.1687 & -0.3313 & 0.5000 & 128 \\ 0.5000 & -0.4187 & -0.0813 & 128 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1.40200 & 0 \\ 1 & -0.34414 & -0.71414 \\ 1 & 1.77200 & 0 \end{bmatrix} \begin{bmatrix} Y \\ Cb-128 \\ Cr-128 \end{bmatrix}$$

彩色图像处理



- 彩色基础知识
- 彩色空间
- 伪彩色处理
- 全彩色图像处理
- 彩色变换
- 彩色图像平滑和尖锐化

伪彩色图像处理



- 什么叫伪彩色图像处理?
  - 也叫假彩色图像处理
  - 根据一定的准则对灰度值赋以彩色的处理
  - 区分：伪彩色图像、真彩色图像、单色图像
- 为什么需要伪彩色图像处理?
  - 人类可以辨别上千种颜色和强度
  - 只能辨别二十几种灰度

伪彩色图像处理



- 应用
  - ✓ 为人们观察和解释图像中的灰度目标

- 怎样进行伪彩色图像处理？
  1. 强度分层技术
  2. 灰度级到彩色转换技术

1、强度分层技术



- 把一幅图像描述为三维函数（x, y, f(x, y)）
- 分层技术：放置平行于(x, y)坐标面的平面
- 每一个平面在相交区域切割图像函数

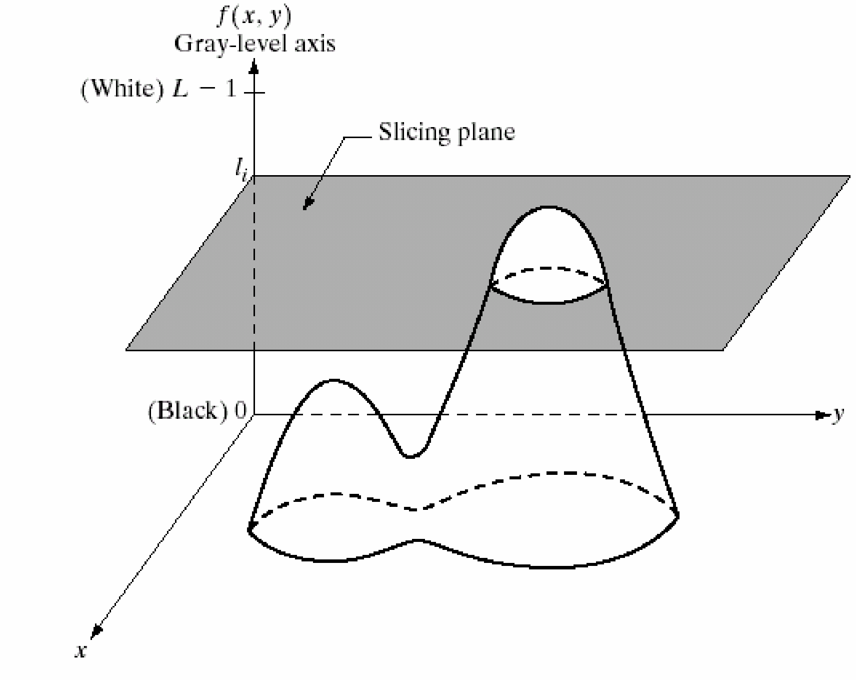


FIGURE 6.18 Geometric interpretation of the intensity-slicing technique.

1、强度分层技术——定义



令[0, L-1]表示灰度级，使l_0代表黑色（f(x, y)=0），l_{L-1}代表白色(f(x, y)=L-1)。假设垂直于强度轴的P个平面定义为量级$l_1, l_2, \ldots, l_P$。0<P<L-1，P个平面将灰度级分为P+1个间隔，$V_1, V_2, \ldots, V_{P+1}$，则灰度级到彩色的赋值关系：

$$f(x,y) = c_k \qquad f(x,y) \in V_k$$

c_k 是与强度间隔 V_k 第K级强度有关的颜色

V_k 是由在l=k-1和l=k分割平面定义的

1、强度分层技术



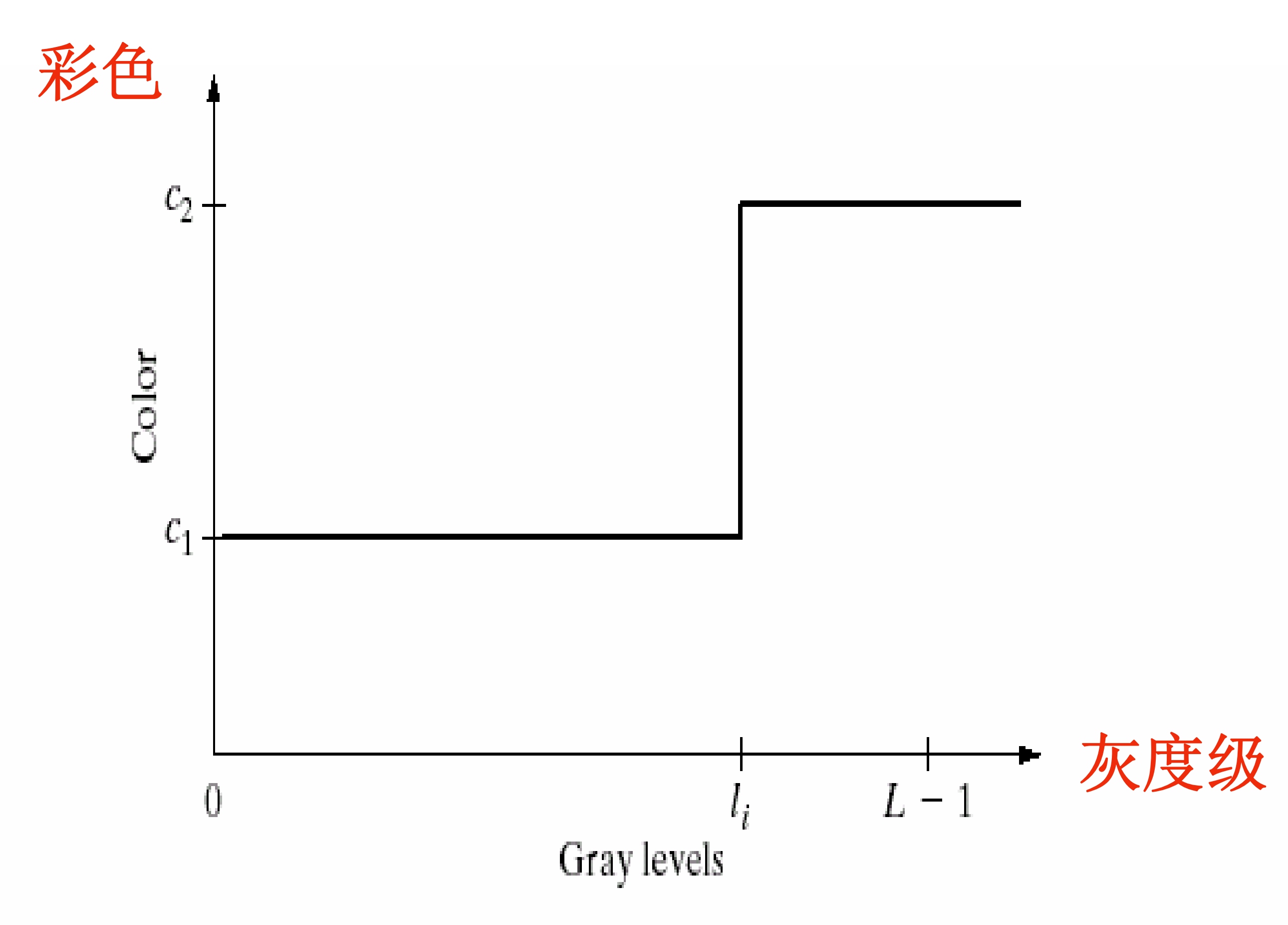
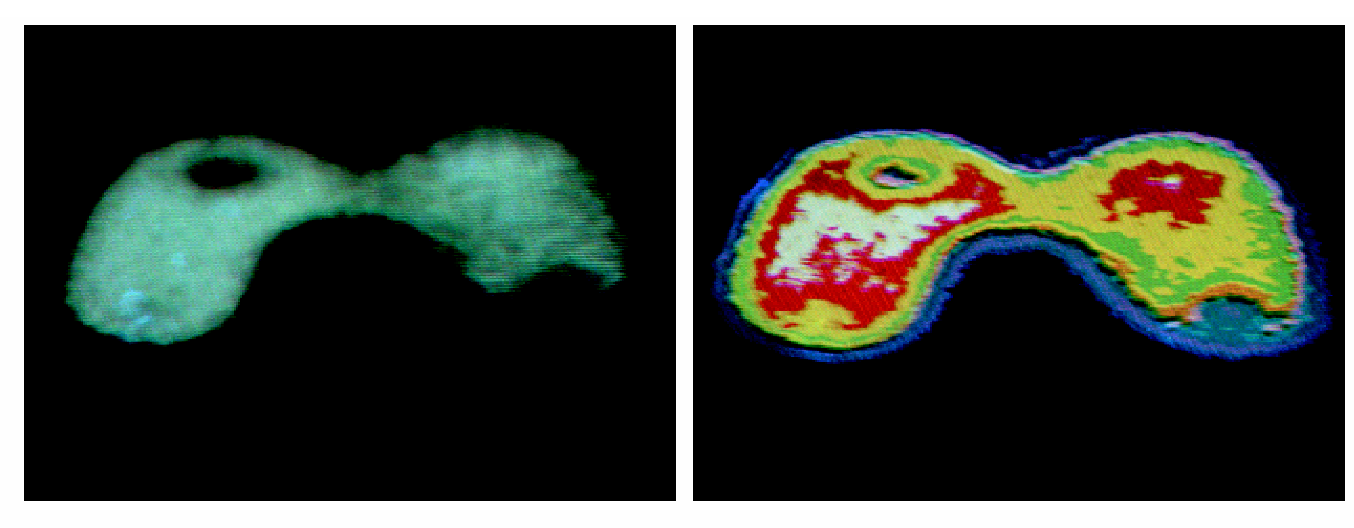


FIGURE 6.19 An alternative representation of the intensity-slicing technique.

例1：甲状腺模型



单色图像　　强度分层结果，8个彩色区域



a b

FIGURE 6.20 (a) Monochrome image of the Picker Thyroid Phantom. (b) Result of density slicing into eight colors. (Courtesy of Dr. J. L. Blankenship, Instrumentation and Controls Division, Oak Ridge National Laboratory.)

左图的恒定强度难以区分病变，右图强度分层结果，清楚的显示恒定强度的不同区域



例2：焊点问题检测

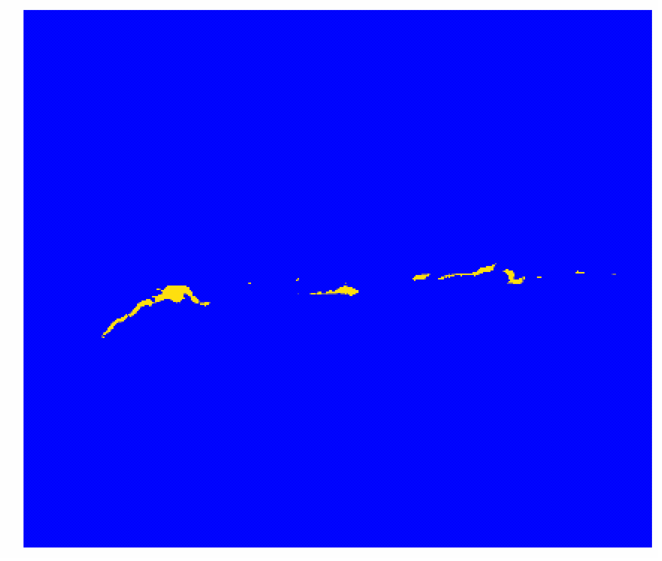
- 图像灰度为255，焊点有问题



a
b

FIGURE 6.21 (a) Monochrome X-ray image of a weld. (b) Result of color coding. (Original image courtesy of X-TEK Systems, Ltd.)





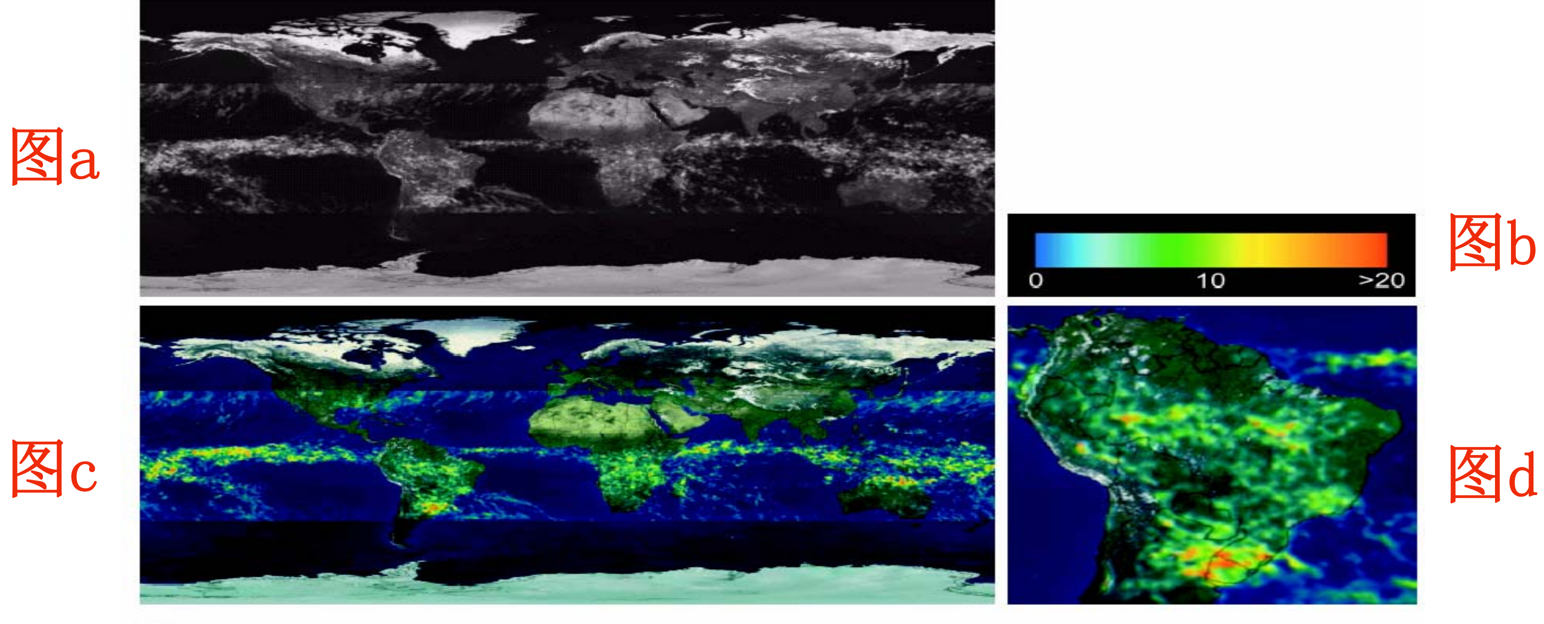
- 给255灰度赋以一种颜色，其它为另一种颜色

- 简化工作，降低误识率



例3：用颜色突出降雨水平

- 图a：图像的强度值直接与降雨相对应，目测困难
- 图b：蓝色表示低降雨量，红色表示高降雨量
- 图c和图d更加清楚



a b
c d

FIGURE 6.22 (a) Gray-scale image in which intensity (in the lighter horizontal band shown) corresponds to average monthly rainfall. (b) Colors assigned to intensity values. (c) Color-coded image. (d) Zoom of the South America region. (Courtesy of NASA.)

2、灰度级到彩色的转换



- 对任何输入像素的灰度级执行3个独立变换
- 3个变换结果分别送入彩色监视器的红、绿、蓝三个通道
- 产生一幅合成图像

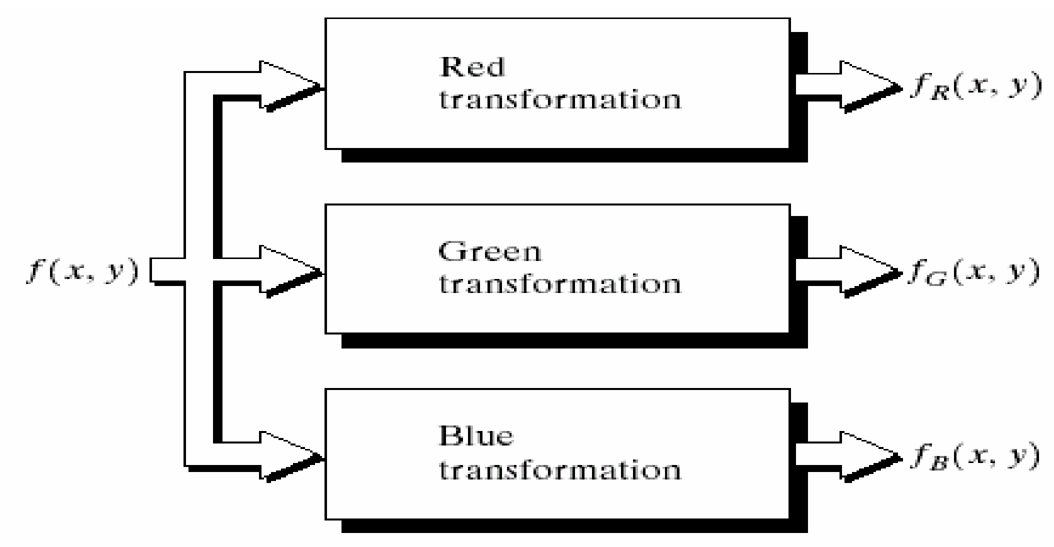


FIGURE 6.23 Functional block diagram for pseudocolor image processing. f_R, f_G, and f_B are fed into the corresponding red, green, and blue inputs of an RGB color monitor.

彩色图像处理



- 彩色基础知识
- 彩色空间
- 伪彩色处理
- 全彩色图像处理
- 彩色变换
- 彩色图像平滑和尖锐化

全彩色图像处理基础



- 全彩色图像处理研究分为两大类：
  - ✓ 分别处理每一分量图像，然后，合成彩色图像
  - ✓ 直接对彩色像素处理：3个颜色分量表示像素向量。令c代表RGB彩色空间中的任意向量

$$c = \begin{bmatrix} c_R \\ c_G \\ c_B \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

全彩色图像处理基础



对大小为 $M \times N$ 的图像

$$c(x,y)=\begin{bmatrix} c_R(x,y) \\ c_G(x,y) \\ c_B(x,y) \end{bmatrix}=\begin{bmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{bmatrix}$$

$$x=0,1,2,\ldots,M-1$$

$$y=0,1,2,\ldots,N-1$$

彩色图像处理



- 彩色基础知识
- 彩色空间
- 伪彩色处理
- 全彩色图像处理
- 彩色变换
- 彩色图像平滑和尖锐化

彩色变换



- 彩色变换函数

$$g(x,y)=T[f(x,y)]$$

$f(x,y)$ 是彩色输入图像

$g(x,y)$ 是变换或处理过的彩色输出图像

T 是在空间邻域 (x,y) 上对 f 的操作

彩色变换



- 彩色变换的简单形式

$$s_i = T_i(r_1, r_2, \ldots, r_n) \qquad i = 1, 2, \ldots, n$$

r_i 和 s_i是 $f(x, y)$ 和$g(x, y)$ 在任何点处彩色分量的变量

$\{T_1, T_2, \ldots T_n\}$ 是一个对 r_i 操作产生 s_i 的变换或彩色映射函数集

选择的彩色空间决定n的值，如RGB彩色空间，n=3，r_1, r_2和r_3表示红、绿、蓝分量;CMYK，则n=4



彩色变换



Full color

草莓和咖啡杯的彩色图像

青　　深红√　黄√　黑

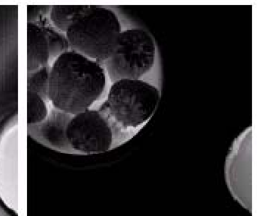


Cyan

Magenta



Yellow



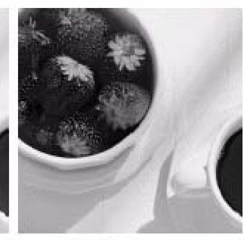
Black

CMYK图像，可知草莓由深红和黄组成

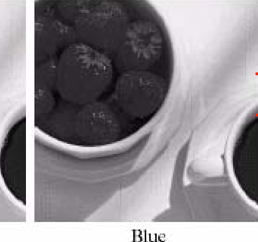
红√



Red



Green



Blue

RGB图像，红色多而蓝色和绿色少

Hue





Saturation



Intensity

HSI图像

补色



补色：在如图所示的彩色环上，与一种色调直接相对立的另一种色调称为补色

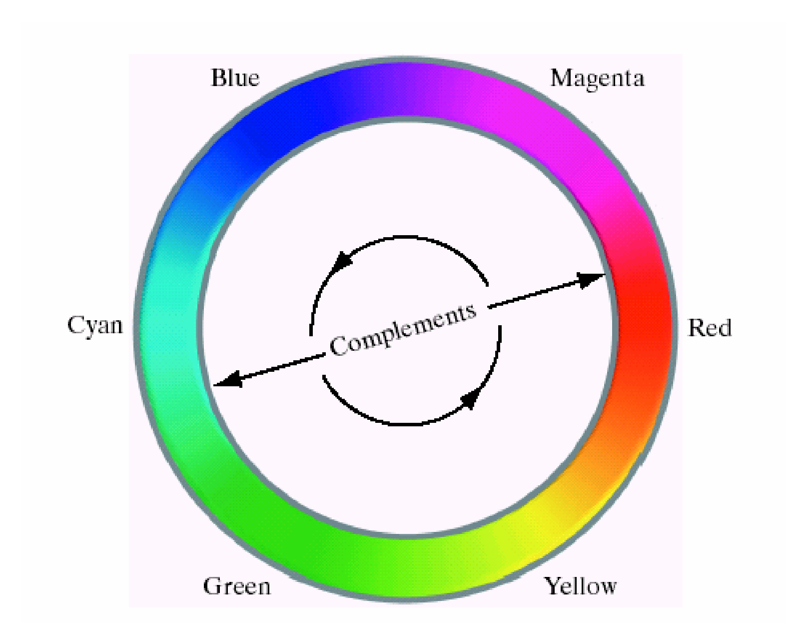


FIGURE 6.32 Complements on the color circle.

作用：增强嵌在彩色图像暗区的细节

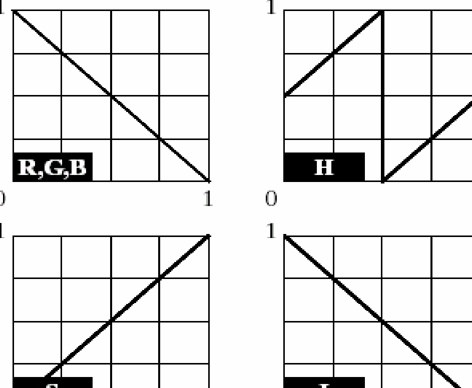
补色

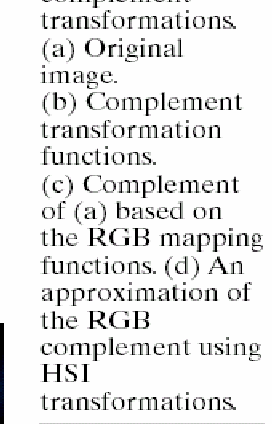


红色在补色中用青色代替，黑色在补色中用白色代替

原图



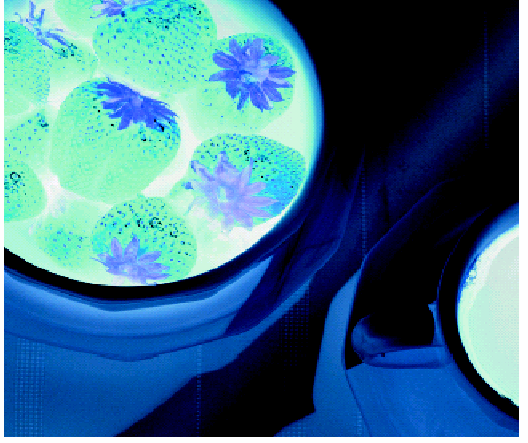


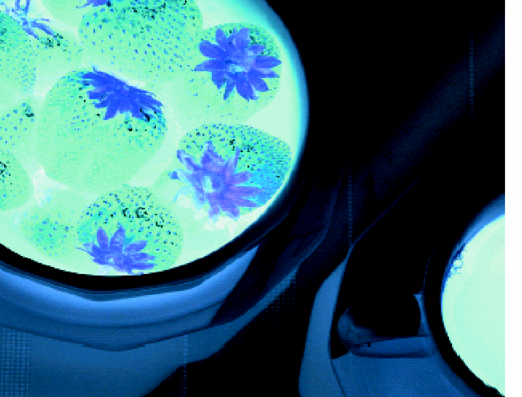


a b
c d

FIGURE 6.33 Color complement transformations. (a) Original image. (b) Complement transformation functions. (c) Complement of (a) based on the RGB mapping functions. (d) An approximation of the RGB complement using HSI transformations.

RGB 补图





HSI 补图

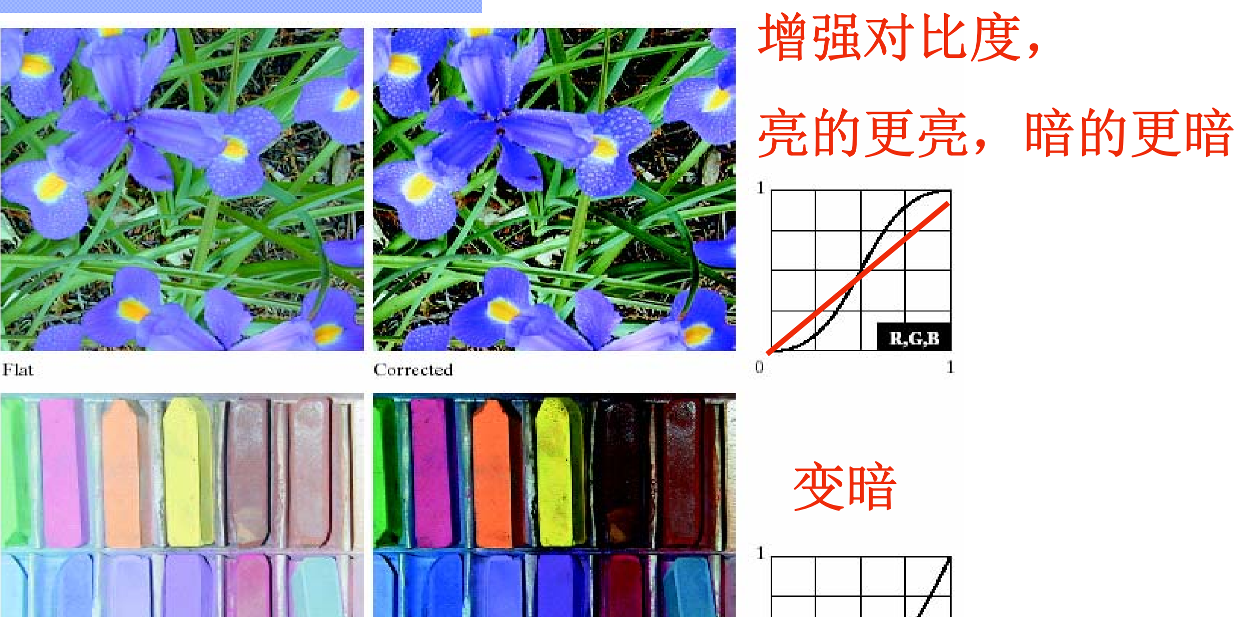
灰度变换



增强对比度，

亮的更亮，暗的更暗

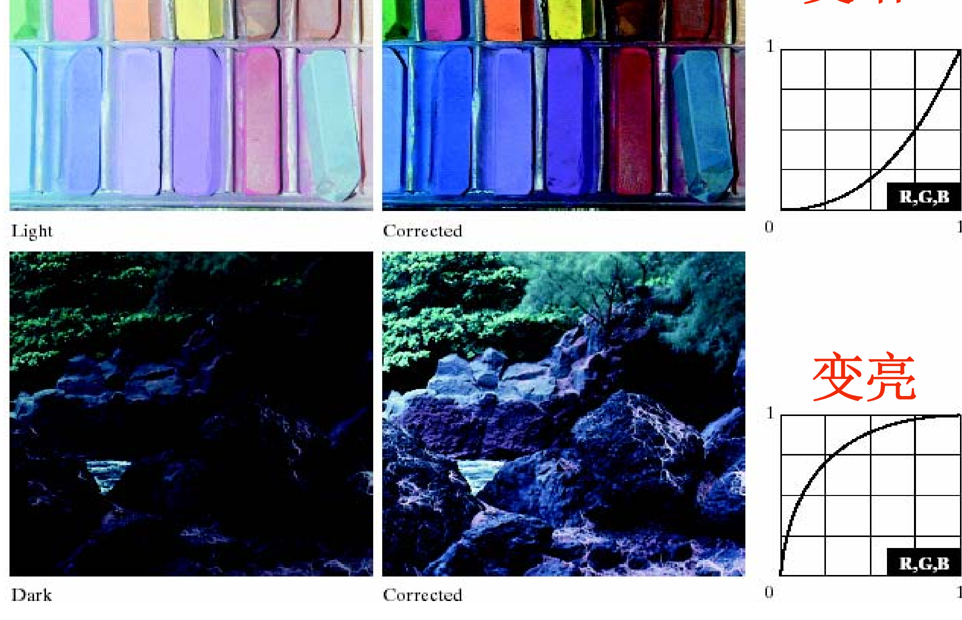
平淡



变暗

亮

暗



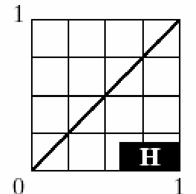
变亮

HSI颜色空间下的直方图均衡化

$$s_k = T(r_k) = \sum_{j=0}^{k} p_r(r_j) = \sum_{j=0}^{k} \frac{n_j}{n}$$

原图





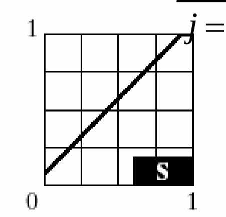




FIGURE 6.37 Histogram equalization (followed by saturation adjustment) in the HSI color space.

1
0.5
0
0.36
1
I



处理前，大量暗彩色

Histogram after processing (median = 0.5)

处理后的直方图

均衡强度

分量的结果





增加图像的饱和度

因为强度改变影响

图像的彩色状态

彩色图像处理



- 彩色基础知识
- 彩色空间
- 伪彩色处理
- 全彩色图像处理
- 彩色变换
- 彩色图像平滑和尖锐化

彩色图像平滑



令S_{xy}表示在RGB彩色图像中定义一个中心在(x, y)的邻域的坐标集，在该邻域中RGB分量的平均值为

$$\bar{c}(x,y)=\frac{1}{k}\sum_{(x,y)\in s_{xy}} c(x,y)$$

$$\bar{c}(x,y)=\begin{bmatrix} \frac{1}{K}\sum_{(x,y)\in s_{xy}} R(x,y) \\ \frac{1}{K}\sum_{(x,y)\in s_{xy}} G(x,y) \\ \frac{1}{K}\sum_{(x,y)\in s_{xy}} B(x,y) \end{bmatrix}$$

彩色图像平滑（5×5 平均模板）



原始彩色图像





红分量图像

a	b
c	d

FIGURE 6.38
(a) RGB image.
(b) Red component image.
(c) Green component.
(d) Blue component.

绿分量图像



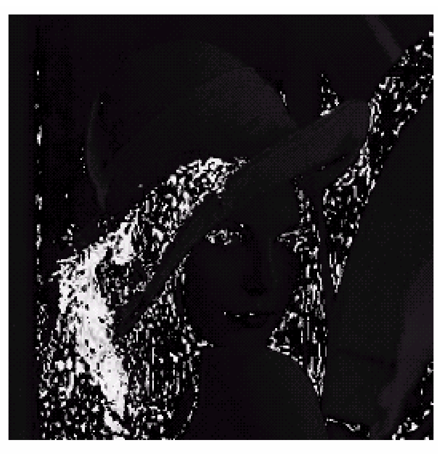


蓝分量图像

彩色图像平滑



H分量图像　　S分量图像　　I分量图像







a b c

FIGURE 6.39 HSI components of the RGB color image in Fig. 6.38(a). (a) Hue. (b) Saturation. (c) Intensity.

彩色图像平滑（HSI图像仅仅处理强度分量）



5×5的灰度平均模板平滑结果

RGB图像平滑结果

平滑强度分量，并转换到RGB图像显示

a图像和b图像的差别图像



a b c　图a　图b　图c

FIGURE 6.40 Image smoothing with a 5 × 5 averaging mask. (a) Result of processing each RGB component image. (b) Result of processing the intensity component of the HSI image and converting to RGB. (c) Difference between the two results.

图c的原因：图a像素的平均是不同彩色的平均，而图b仅仅是强度的平均，原彩色（色调H和饱和度S）保持不变

彩色图像尖锐化（拉普拉斯微分）



RGB彩色空间，分别计算每一分量图像的拉普拉斯变换

$$\nabla^2\left[c(x,y)\right]=\begin{bmatrix}\nabla^2 R(x,y)\\ \nabla^2 G(x,y)\\ \nabla^2 B(x,y)\end{bmatrix}$$

$$g(x,y)=f(x,y)-\nabla^2 f(x,y)$$

$$=f(x,y)-\{[f(x+1,y)+f(x-1,y+f(x,y+1)+f(x,y-1))]-4f(x,y)\}$$

$$=5f(x,y)-[f(x+1,y)+f(x-1,y)+f(x,y+1)+f(x,y-1)]$$

彩色图像尖锐化（拉普拉斯微分）



图6.38RGB图像的拉普拉斯变换	图6.39HSI图像的强度I分量图像的拉普拉斯变换	a图像和b图像的差别图像

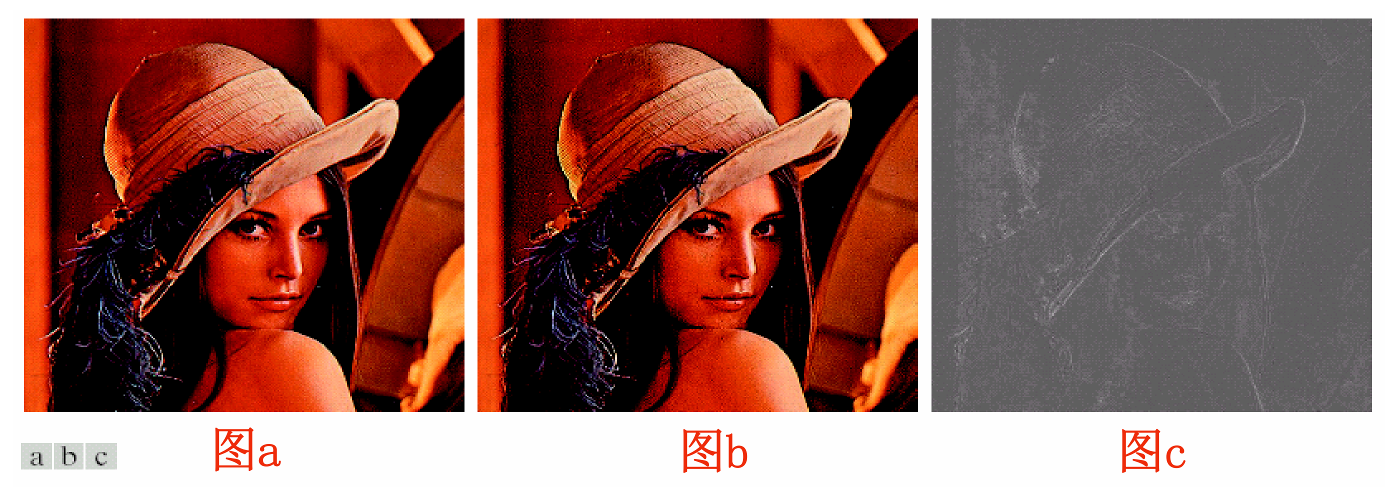


FIGURE 6.41 Image sharpening with the Laplacian. (a) Result of processing each RGB channel. (b) Result of processing the intensity component and converting to RGB. (c) Difference between the two results.

图c的原因：图a像素的锐化是不同彩色的锐化，而图b仅仅是强度的锐化，原彩色（色调H和饱和度S）保持不变

彩色分割（把一幅图像分成区域）



- HSI彩色空间分割——直观
  - ✓ H色调图像方便描述彩色
  - ✓ S饱和度图像做模板分离感兴趣的特征区
  - ✓ I强度图像不携带彩色信息

- RGB彩色空间——直接

HSI彩色空间分割



阈值产生的二值图像：

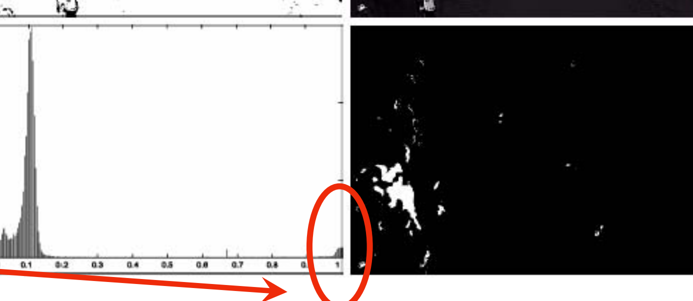
饱和度图像中阈值等于

最大饱和度的10%，大于阈值的像素赋1，其它赋0









RGB彩色空间分割



令z代表RGB空间中的任意一点，a是分割颜色样本集的平均颜色向量

$$D(z,a)=\|z-a\|=\sqrt{(z_R-a_R)^2+(z_G-a_G)^2+(z_B-a_B)^2}$$

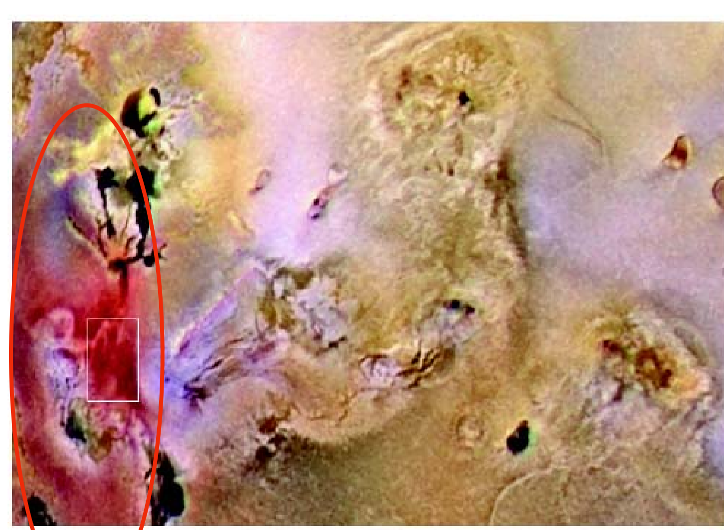
D_0是距离阈值

如果 $D(z,a)\le D_0$，则z和a相似

如果 $D(z,a)>D_0$ 则z和a不相似

RGB彩色空间分割





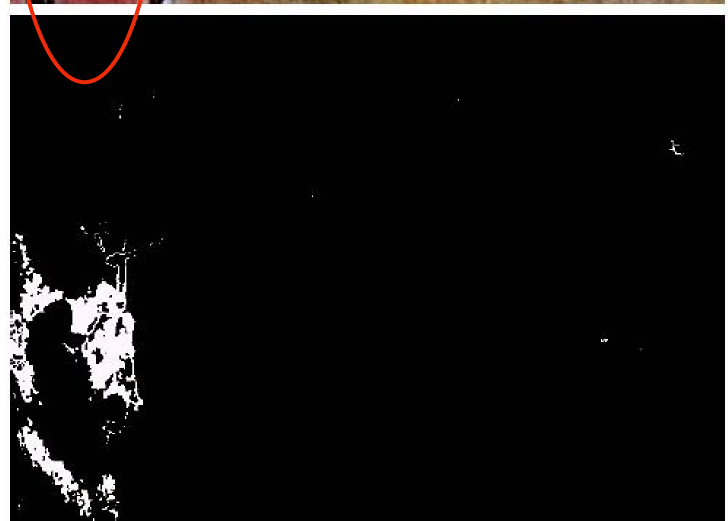
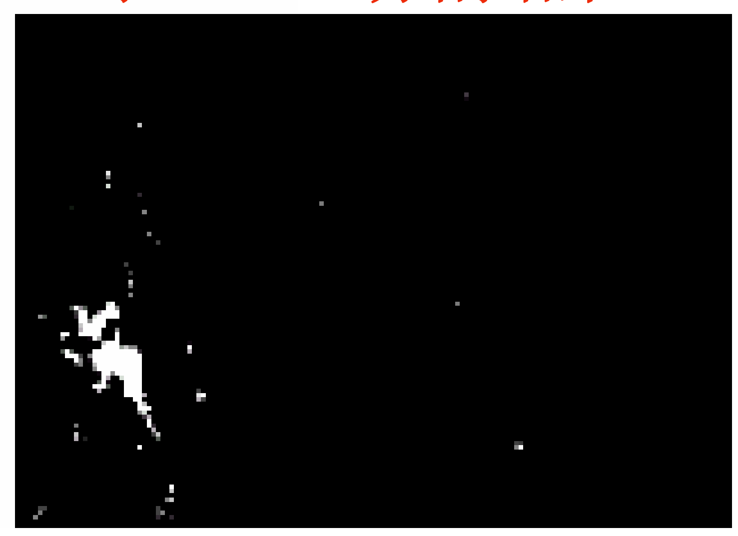




FIGURE 6.44 Segmentation in RGB space. (a) Original image with colors of interest shown enclosed by a rectangle. (b) Result of segmentation in RGB vector space. Compare with Fig. 6.42(h).

对比：HSI分割结果



数字图像处理（3）



任何问题？

研究生课程



数字图像处理

Digital Image Processing

彭宇新

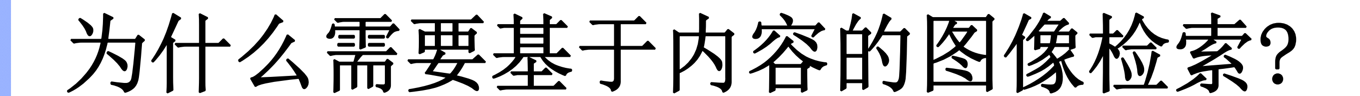
北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

基于内容的图像检索



- 为什么需要基于内容的图像检索?
- 查询方式，查询demo，现有系统简介
- 具体内容
  - ✓ 特征提取
  - ✓ 相似度匹配
  - ✓ 相关反馈
  - ✓ 索引结构
- MPEG-7介绍：性能评价等
- 思考的几个问题?

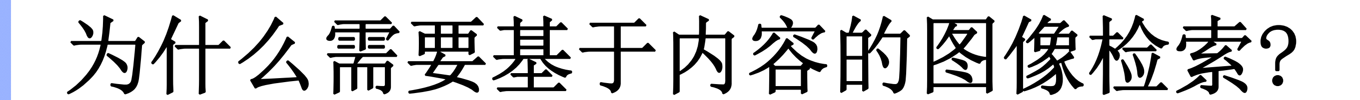


为什么需要基于内容的图像检索?



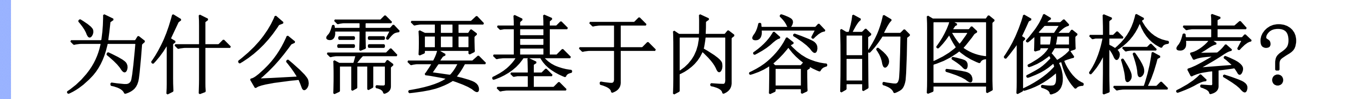
- 当前图像内容的特征
  - ✓ 海量的图像内容出现
    - ➢ 人类已有的：历史、地理、军事、医学…
    - ➢ 每天新增的：数码相机、互联网
  - ✓ 图像内容的特征：没有索引、目录或摘要
  - ✓ 怎样找到需要的图像？

为什么需要基于内容的图像检索?





- 基于文字描述的图像检索
  - ✓ 丰富的图像内容很难用文字来全面描述
  - ✓ 文字的选取因人而异，带有很大的主观性
  - ✓ 耗费大量的人力和时间
- 结论：对于海量的图像，基于人工标注文字的检索已不可能
- 问题：目前搜索引擎提供的基于文字的图像检索，能够解决问题吗？解决得好吗？





- 为什么需要基于内容的图像检索(CBIR)?
  - ✓提供图像的检索功能
  - ✓不需要人手工标注，计算机自动分析和检索，节省人力和时间
  - ✓使海量图像的管理和索引成为可能
  - ✓存在的问题：
    - ➢ 人类高级语义特征与计算机底层特征表示的矛盾
    - ➢ 查询方式问题

基于内容的图像检索

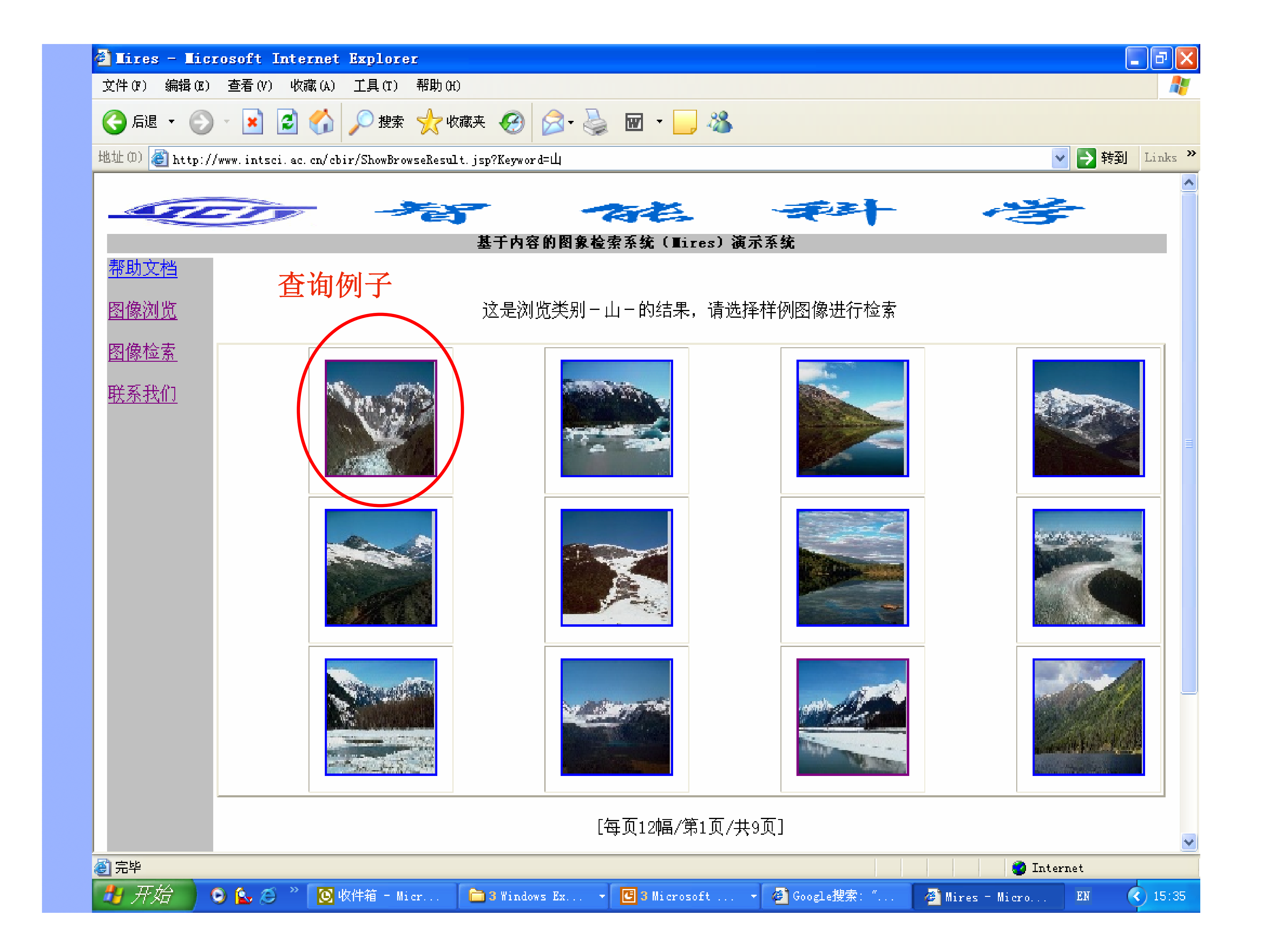


- 为什么需要基于内容的图像检索?
- 查询方式，查询demo，现有系统简介
- 具体内容
  - ✓ 特征提取
  - ✓ 相似度匹配
  - ✓ 相关反馈
  - ✓ 索引结构
- MPEG-7介绍：性能评价等
- 思考的几个问题？

基于内容的图像检索



- 查询方式
  - ✓ 利用采样图像：选择图像数据库中的图像
  - ✓ 利用范例图像：来自图像库之外，用户自己提供
  - ✓ 利用局部图像：先提取图像区域，再利用该图像区域检索；或利用不同图像的不同区域的拼图
  - ✓ 利用绘制图像：用户绘制出来的图作为查询图像，描述用户的检索目标或场景
    - ➢ 草图：简单的素描，用户期望的目标形状
    - ➢ 绘画：用颜色描绘图像区域及区域的空间分布
  - ✓ ……


Mires - Microsoft Internet Explorer
文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)
后退 搜索 收藏夹
地址(D) http://www.intsci.ac.cn/cbir/ShowBrowseResult.jsp?Keyword=山
转到 Links
基于内容的图象检索系统（Mires）演示系统
帮助文档
图像浏览
图像检索
联系我们
查询例子
这是浏览类别－山－的结果，请选择样例图像进行检索
[每页12幅/第1页/共9页]
完毕
Internet
开始
收件箱 - Micr...
3 Windows Ex...
3 Microsoft ...
Google搜索："...
Mires - Micro...
EN
15:35

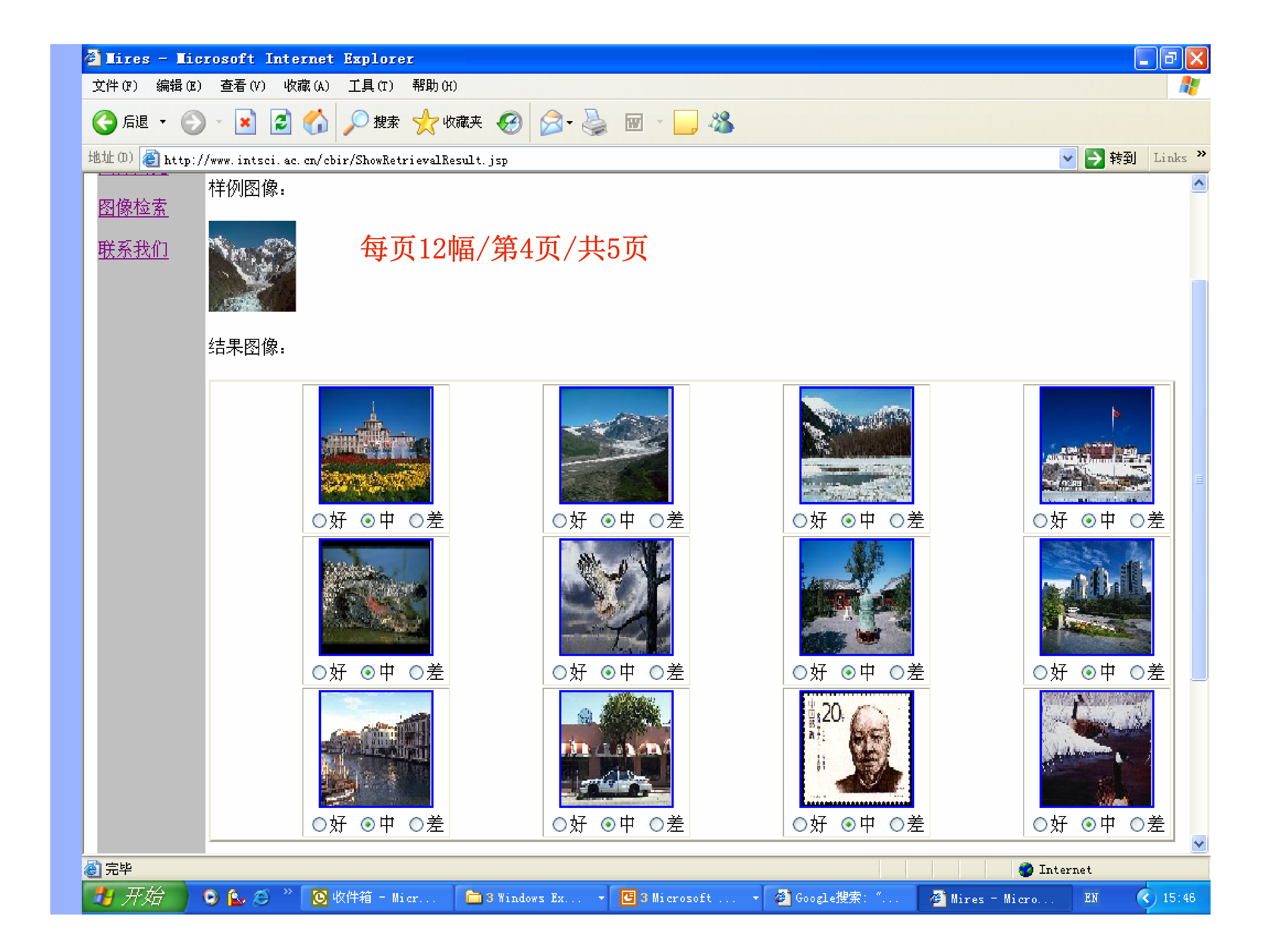

Mires - Microsoft Internet Explorer
文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)
后退 搜索 收藏夹
地址(D) http://www.intsci.ac.cn/cbir/RetrievalSetup.jsp?ImageLoad=Image\mountain\114004.jpg
转到 Links
基于内容的图象检索系统(Mires)演示系统
帮助文档
图像浏览
图像检索
联系我们
进行检索的样例图像：
颜色权值 1 纹理权值 1 形状权值 1
请选择检索的关键字： 全部
检索 取消
中国科学院计算技术研究所智能信息处理重点实验室 版权所有
完毕
Internet
开始
收件箱 - Micr...
3 Windows Ex...
3 Microsoft ...
Google搜索："...
Mires - Micro...
EN
15:39

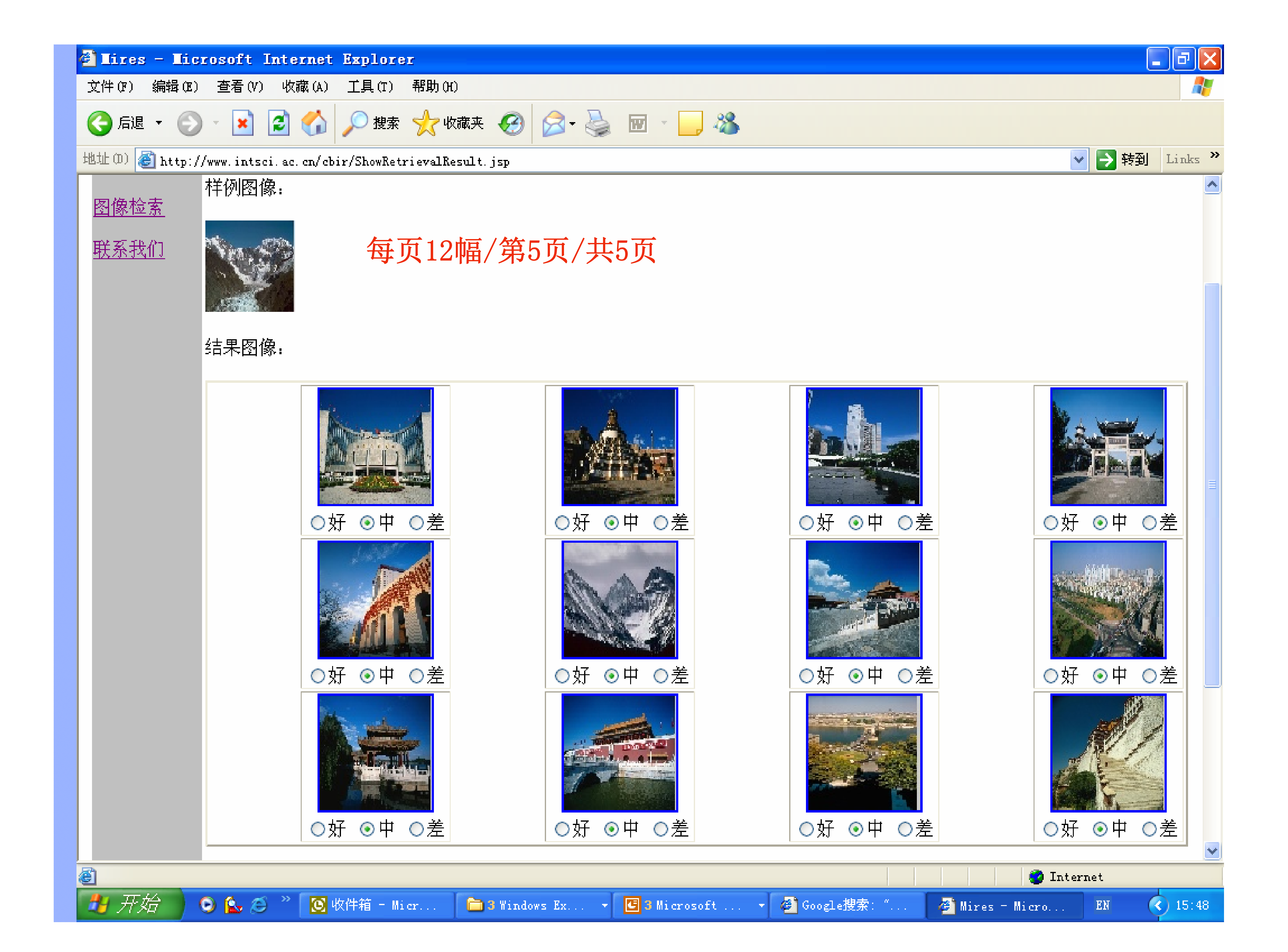

Mires - Microsoft Internet Explorer
文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)
后退 搜索 收藏夹
地址(D) http://www.intsci.ac.cn/cbir/ShowRetrievalResult.jsp
转到 Links
图像检索
联系我们
样例图像：
每页12幅/第1页/共5页
结果图像：
好 中 差
Internet
开始
15:41


Mires - Microsoft Internet Explorer
文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)
后退 搜索 收藏夹
地址(D) http://www.intsci.ac.cn/cbir/ShowRetrievalResult.jsp
转到 Links
图像检索
联系我们
样例图像：
每页12幅/第2页/共5页
结果图像：
好 中 差
Internet
开始
15:43



Mires - Microsoft Internet Explorer
文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)
后退 搜索 收藏夹
地址(D) http://www.intsci.ac.cn/cbir/ShowRetrievalResult.jsp
转到 Links
图像检索
联系我们
样例图像：
每页12幅/第3页/共5页
结果图像：
好 中 差
Internet
开始
15:45


Mires - Microsoft Internet Explorer
文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)
后退 搜索 收藏夹
地址(D) http://www.intsci.ac.cn/cbir/ShowRetrievalResult.jsp
转到 Links
图像检索
联系我们
样例图像：
每页12幅/第4页/共5页
结果图像：
好 中 差
完毕
Internet
开始
收件箱 - Micr...
3 Windows Ex...
3 Microsoft ...
Google搜索："...
Mires - Micro...
EN
15:46


Mires - Microsoft Internet Explorer
文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)
后退 搜索 收藏夹
地址(D) http://www.intsci.ac.cn/cbir/ShowRetrievalResult.jsp
转到 Links
图像检索
联系我们
样例图像:
每页12幅/第5页/共5页
结果图像:
好 中 差
Internet
开始
收件箱 - Micr...
3 Windows Ex...
3 Microsoft ...
Google搜索: "...
Mires - Micro...
EN
15:48

基于内容的图像检索



- 现有系统
  - QBIC
  - Virage
  - Photobook
  - VisualSEEK
  - MARS
  - …

基于内容的图像检索



- QBIC(Query By Image Content)
  - ✓ IBM研制开发
  - ✓ 支持图像例子、绘制草图或定制图像特征模板来检索图像库
  - ✓ 颜色使用RGB,YIQ,Lab和MTM颜色空间
  - ✓ 纹理使用改进的Tamura表示：粗糙度，对比度和方向性等
  - ✓ 形状特征有目标面积、各阶矩、离心率和主轴方向等
  - ✓ 草图特征，如边界各点的切线方向等
  - ✓ 开发了基于聚类的索引技术

基于内容的图像检索



- Virage的VIR Image Engine
  - ✓ 类似于QBIC系统
  - ✓ 特色：将特征进行归类，用户查询时可以对特征进行任意组合
- Photobook
  - ✓ MIT媒体实验室开发
  - ✓ 首先根据目标类别划分数据库，3个基本目标是：脸、形状和纹理

基于内容的图像检索



- 为什么需要基于内容的图像检索?
- 查询方式，查询demo，现有系统简介
- 具体内容
  - ✓ 特征提取
  - ✓ 相似度匹配
  - ✓ 相关反馈
  - ✓ 索引结构
- MPEG-7介绍：性能评价等
- 思考的几个问题?

基于内容的图像检索



- 特征提取(Feature Extraction)
  - 颜色特征
  - 纹理特征
  - 形状特征
  - 空间关系特征
  - ……

基于内容的图像检索



- 颜色特征
  - ✓ 颜色空间（也称彩色模型或彩色系统）
  - ✓ 颜色特征表示
  - ✓ 颜色相似度匹配算法

颜色特征



- 颜色空间
  - ✓ RGB
  - ✓ CMY或CMYK
  - ✓ HSI
  - ✓ HSV:比HSI模型更与人类对颜色的感知接近
  - ✓ 均匀颜色空间模型：MTM，LUV和Lab

颜色空间



- RGB ⟶ HSV

$$H=\begin{cases}\arccos\dfrac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} & B\le G\\[2ex] 2\pi-\arccos\dfrac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} & B>G\end{cases}$$

$$S=\frac{\max(R,G,B)-\min(R,G,B)}{\max(R,G,B)}$$

$$V=\frac{\max(R,G,B)}{255}$$

颜色特征



- 均匀颜色空间模型
  - ✓ 从图像处理的角度，对颜色的描述应该与人对颜色的感知越接近越好
  - ✓ 从视觉感知均匀的角度，人所感知到的两个颜色的距离应该与这两个颜色在表达它们的颜色空间中的距离越成比例越好
  - ✓ 如果在一个颜色空间中，人所观察到的两种颜色的区别程度与该颜色空间中两点间的欧式距离对应，则称该空间为均匀颜色空间
  - ✓ 典型的例子：MTM, Luv和Lab

颜色特征

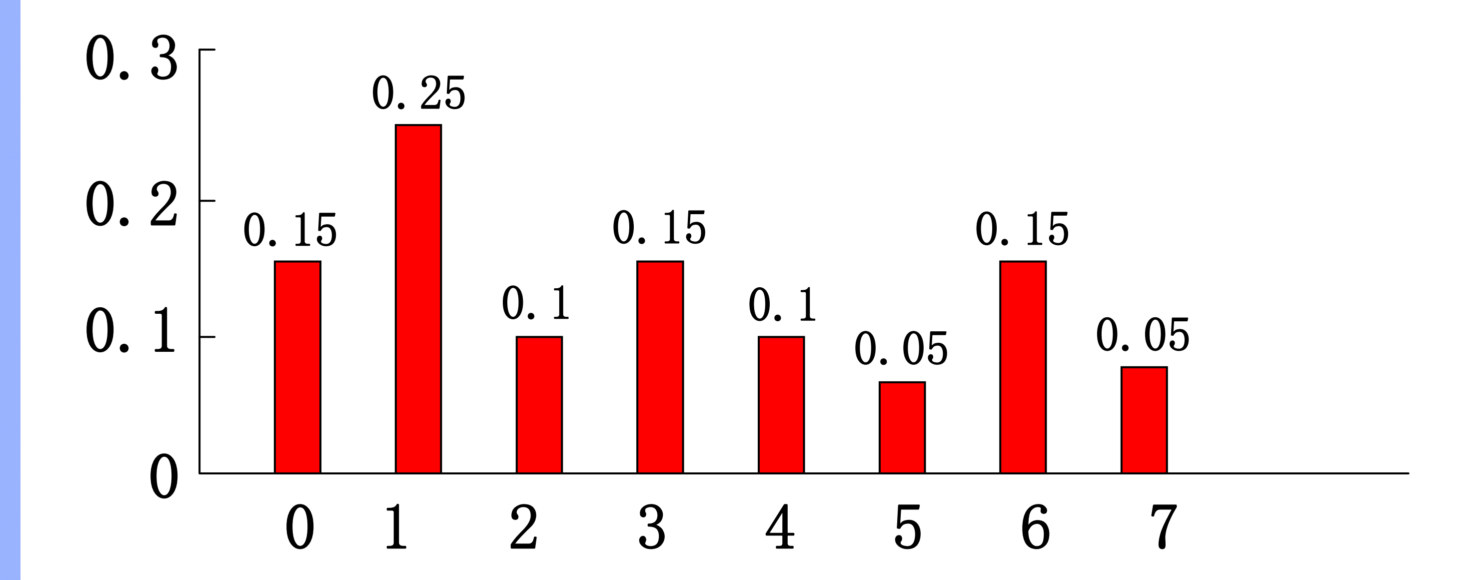


- 颜色特征表示
  - ✓ 统计直方图
  - ✓ 累积直方图
  - ✓ 颜色布局
  - ✓ 颜色分块



颜色特征

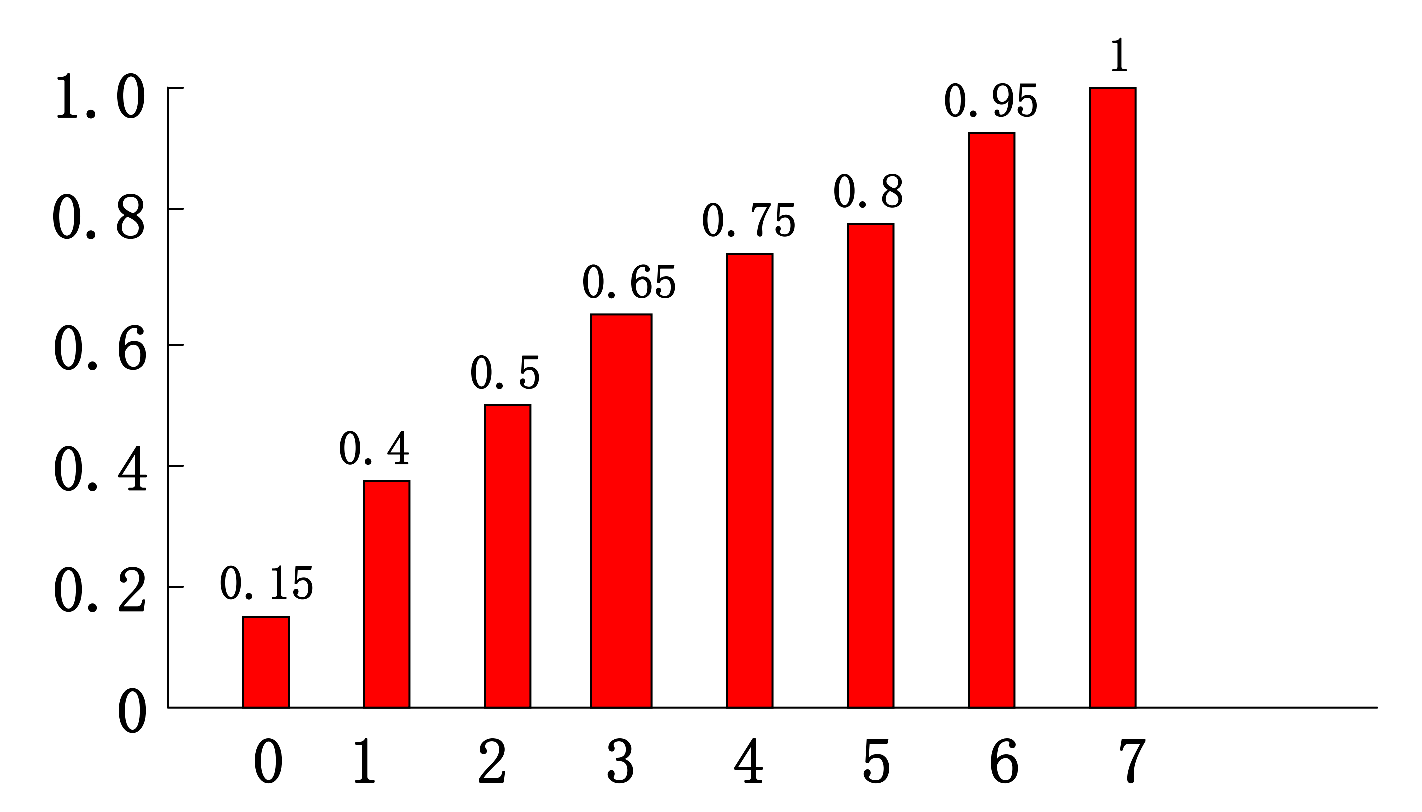
- 统计直方图：$p(k)=\frac{n_k}{n} \quad k=0,1,\ldots,L-1$



颜色特征



- 累积直方图：$I(k)=\sum_{i=0}^{k}\frac{n_k}{n} \quad k=0,1,\ldots,L-1$



颜色特征



- 颜色布局
  - ✓ MPEG-7中建议了一种颜色布局描述符(color layout)
  - ✓ 表达颜色的空间分布信息
- 颜色布局算法
  - ✓ 将图像从RGB空间映射到YCbCr空间：

$$Y = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

$$Cb = -0.169 \times R - 0.331 \times G + 0.500 \times B$$

$$Cr = 0.500 \times R - 0.419 \times G - 0.081 \times B$$

颜色特征



- 颜色布局算法
  - ✓ 将整副图像平均分成64块，计算每一块中所有像素各颜色分量的平均值，以此作为该块的代表颜色（主颜色）
  - ✓ 将各块的平均值数据进行DCT变换
  - ✓ 通过之字形扫描和量化，取出3组颜色DCT变换后的低频分量，构成该图像的颜色布局描述符

颜色特征



- 颜色相似度匹配算法

1. 直方图相交法
2. 欧式距离
3. 距离法
4. 中心矩法
5. 参考颜色表法

颜色特征



1、直方图相交法

令 $H_Q(k)$ 和 $H_D(k)$ 分别为查询图像Q和数据库图像D的（某一）特征的统计直方图，则两图像之间的相似值为

$$P(Q,D)=\frac{\sum_{k=0}^{L-1}\min\{H_Q(k),H_D(k)\}}{\sum_{k=0}^{L-1}H_Q(k)}$$

距离为： $d_{L_1}=1-P(Q,D)$

颜色特征



1、直方图相交法

如果用HSV直方图表示每幅图像，则两幅图像Q和D的直方图交表示为：

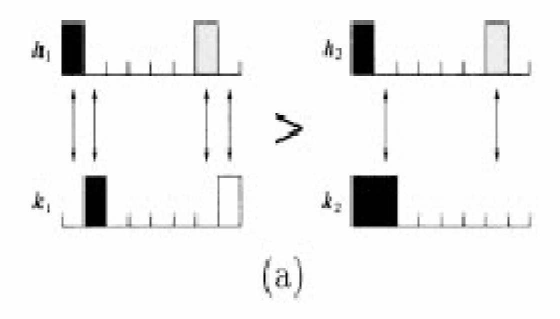
$$P(Q,D)=\frac{\sum_{H}\sum_{S}\sum_{V}\min\left\{H_Q(H,S,V),H_D(H,S,V)\right\}}{\min\left\{\sum_{H}\sum_{S}\sum_{V}H_Q(H,S,V),\sum_{H}\sum_{S}\sum_{V}H_D(H,S,V),\right\}}$$

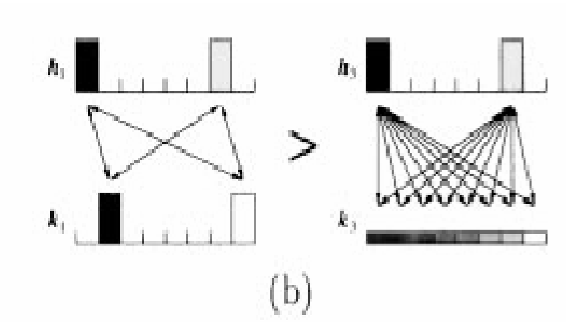
$$P(Q,D)=\frac{\sum_{H}\sum_{S}\sum_{V}\min\left\{H_Q(H,S,V),H_D(H,S,V)\right\}}{\sum_{H}\sum_{S}\sum_{V}H_Q(H,S,V)}$$

直方图相交法的缺点：思考题

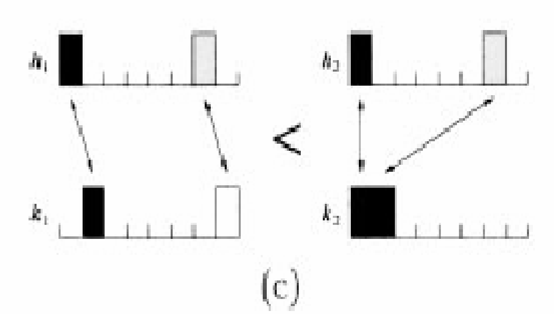


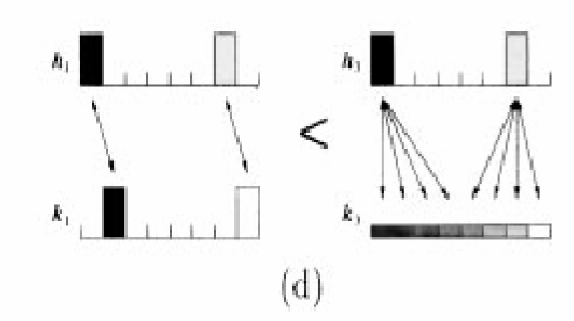
Bin-By-Bin





Bin-Across-Bin





颜色特征



2、欧式距离法

$$P_E(Q,D)=\sqrt{\sum_{i=0}^{L-1}\left[H_Q(i)-H_D(i)\right]^2}$$

加权的欧式距离法

$$P_w(Q,D)=\sqrt{\sum_{i=0}^{L-1}W_i\left[H_Q(i)-H_D(i)\right]^2}$$

$$W_i=\begin{cases}H_Q(i) & H_Q(i), H_D(i)>0\\ 1 & H_Q(i)\text{或 }H_D(i)=0\end{cases}$$

颜色特征



3、距离法

为减少计算量，可借助直方图的均值来粗略地表达颜色信息。如果图像用R，G，B三个分量表示，则图像的特征向量由这三个分量的均值组成

$$f=\left[\mu_R,\mu_G,\mu_B\right]^T$$

此时图像Q和D的相似值为

$$P(Q,D)=\sqrt{\left(f_Q-f_D\right)^2}=\sqrt{\sum_{R,G,B}\left(\mu_Q-\mu_D\right)^2}$$

颜色特征



3、距离法

MPEG-7中建议的颜色布局描述符在匹配时使用了类似的距离公式

$$P(Q,D)=\sqrt{\sum_i W_Y\left(Y_{i,Q}-Y_{i,D}\right)^2}+\sqrt{\sum_i W_{Cr}\left(Cr_{i,Q}-Cr_{i,D}\right)^2}$$

$$+\sqrt{\sum_i W_{Cb}\left(Cb_{i,Q}-Cb_{i,D}\right)^2}$$

W为权重，Y_i, Cr_i, Cb_i为各个分量的第i个DCT系数

颜色特征



4、中心矩法

- 对直方图来说，均值是其零阶矩，更高阶的矩也可用于匹配
- 对一幅灰度级为L的图像，其直方图为H(i),则

该图像的前三阶中心矩分别为

$$M_1 = \frac{1}{L}\sum_{i=0}^{L-1} H(i)$$

$$M_2 = \left[\frac{1}{L}\sum_{i=0}^{L-1} (H(i) - M_1)^2\right]^{\frac{1}{2}}$$

$$M_3 = \left[\frac{1}{L}\sum_{i=0}^{L-1} (H(i) - M_1)^3\right]^{\frac{1}{3}}$$

颜色特征



4、中心矩法

✓ 对彩色图像，用 $M_{i,QR},M_{i,QG},M_{i,QB}$和 $M_{i,DR},M_{i,DG},M_{i,DB}$ 表示图像Q和D的3个分量R, G, B的直方图的i (i≤3)阶中心矩，则它们之间的匹配值为

$$P(Q,D)=\sqrt{W_R\sum_{i=1}^{3}(M_{i,QR}-M_{i,DR})^2+W_G\sum_{i=1}^{3}(M_{i,QG}-M_{i,DG})^2+W_B\sum_{i=1}^{3}(M_{i,QB}-M_{i,DB})^2}$$

如果彩色图像用3个分量H, S, V来表示，也可用类似方法计算

颜色特征



5、参考颜色表法

- ✓ 距离法太粗糙，直方图相交法计算量太大
- ✓ 折衷的方法：将图像颜色用一组参考色表示，这组参考色应能覆盖视觉上可感受到的各种颜色
- ✓ 参考色的数量要比原图的少，可得简化的直方图，所以匹配的特征向量是

$$f = \left[r_1, r_2, \ldots, r_n \right]$$

r_i是第i种颜色出现的频率，n是参考颜色表的尺寸

颜色特征



5、参考颜色表法

加权后的查询图像Q和数据库图像D之间的匹配值为

$$P(Q,D)=W\sqrt{\left(f_Q-f_D\right)^2}=\sqrt{\sum_{i=1}^{n}W_i\left(r_{i,Q}-r_{i,D}\right)^2}$$

$$W_i=\begin{cases}r_{i,Q} & \text{当 } r_{i,Q}>0 \text{ 且 } r_{i,D}>0\\ 1 & \text{当 } r_{i,Q}=0 \text{ 或 } r_{i,D}=0\end{cases}$$

基于内容的图像检索



- 纹理特征
  - ✓ 概述
  - ✓ 纹理表示
  - ✓ 纹理距离度量举例

基于内容的图像检索



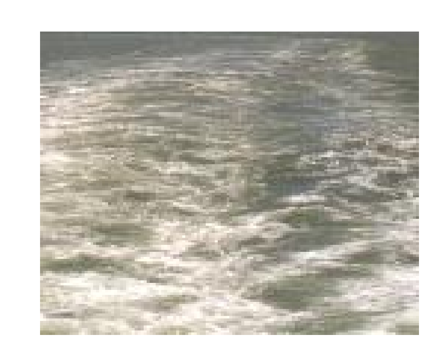
砖墙

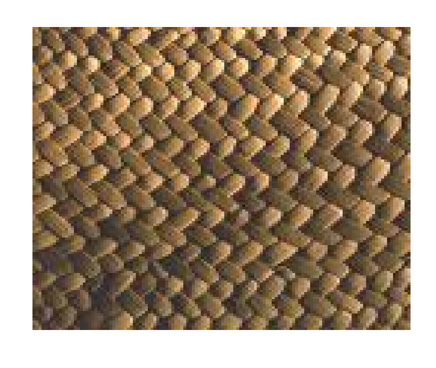


木头



水





布



乱草



树叶

基于内容的图像检索



- 纹理概述
  - ✓ 描述图像时常用的一个概念，类似于颜色，也常取决于感知
  - ✓ 对纹理的感受是与心理效果相结合的，所以用语言或文字来描述纹理很困难
  - ✓ 类似于布纹、草地、砖墙等重复性结构的图像称为纹理图像

基于内容的图像检索



- 纹理概述（续）
  - ✓ 纹理可认为是灰度(颜色)在空间以一定的形式变化而产生的图案（模式）
  - ✓ 一般来说，纹理图像中灰度分布具有某种周期性，即便灰度变化是随机的，它也具有一定的统计特性
  - ✓ 纹理通常和图像频谱中的高频分量是密切联系的，光滑的图像（主要包括低频分量）一般不认为是纹理图像

基于内容的图像检索



- 纹理表示
  - ✓ Tamura表示法
  - ✓ 联合概率矩阵表示
  - ✓ 小波表示
  - ✓ ……

基于内容的图像检索



- Tamura纹理表示法
  - ✓ 对比度、粗细度和方向性是描述纹理的3个量
  - ✓ 利用这3个量可组合成一个3-D空间
  - ✓ 在这个3-D空间中，两点间的欧式距离与人对纹理感知的差距很接近

基于内容的图像检索



- 对比度(contrast)

✓ 依赖于像素的灰度分布，它可测量图像中局部的灰度变化，也常用作对图像整体感知的一种描述

✓ 一般来说，图像的对比度与图像灰度的动态范围及图像中边缘的尖锐程度都有关

$$contrast = \frac{\sigma}{\left[\mu_4 / \sigma^4\right]^{\frac{1}{4}}}$$

其中，σ是图像灰度的标准方差，μ_4 是图像灰度的4阶中心矩

基于内容的图像检索



- 粗细度(contrast，也叫粗糙度)
  - ✓ 测量纹理的间隔尺寸/粒度
  - ✓ 与图像的分辨率有关，分辨率大则纹理比较粗，即组成纹理的元素尺寸比较大

$$coarseness = \frac{1}{mn}\sum_{ij} 2^k f(i,j)$$

其中，窗口尺寸是 $2^k \times 2^k$，k能使X和Y方向的 $\sum_{ij} f(i,j)/2^{2k}$ 最大化，k的取值在0到5之间

基于内容的图像检索



- 方向性(directionality)
  - ✓ 是一个全局的纹理特征，它描述纹理是如何沿某些方向散布或集中的
  - ✓ 一般来说，方向性与纹理基元的形状以及如何将这些纹理基元排列的规则有关

基于内容的图像检索



- 联合概率矩阵表示
  - ✓ 在70年代早期，Haralick等提出了纹理特征的联合概率矩阵表示。这种方法考察了纹理在灰度级空间中的相关性
  - ✓ 首先基于像素间的距离和方向建立联合概率矩阵
  - ✓ 然后从联合概率矩阵中提取出有意义的统计量作为纹理描述

基于内容的图像检索



- 小波表示
  - ✓ Smith和Chang等提出用小波子带的统计量（均值和方差）作为纹理表示
  - ✓ 对112幅Brodatz纹理图像的表示中达到了90%的准确率

一些纹理表示的比较结果



- Weszka等比较了
  - ✓ 傅里叶能量谱
  - ✓ 联合概率矩阵表示
  - ✓ 一次灰度级差分统计量

- 在地形样本上的测试表明傅里叶表示法的性能相对最差

一些纹理表示的比较结果



- Ohanian和Dubes比较了
  - 马尔科夫随机场表示
  - 多通道滤波器表示
  - 基于分形的表示
  - 共生矩阵表示
- 测试：
  - 两个人工纹理测试集（分形）和高斯马尔科夫场
  - 两个自然纹理测试集（羽毛和油漆表面）

一些纹理表示的比较结果



- 结论：共生矩阵是最好的
- Ma和Manjunath比较了
  - ✓ 正交和双正交小波变换表示
  - ✓ 树结构小波变换表示
  - ✓ Gabor小波变换表示
  - ✓ 结论：Gabor小波变换表示效果最好

基于内容的图像检索



- 形状特征
  - ✓ 形状常与目标联系在一起，有一定的语义含义
  - ✓ 要获得有关目标的形状参数，需要先对图像进行分割，所以形状特征的提取会受图像分割效果的影响
  - ✓ 目标形状的描述是一个非常复杂的问题，至今还没有找到形状的确切数学定义

基于内容的图像检索



- 形状特征
  - ✓ 从不同视角获取的图像中目标形状可能会有很大差别
  - ✓ 为准确进行形状匹配，需要解决平移、尺度、旋转变换不变性的问题

基于内容的图像检索



- 通常，形状表示方法可分为两类
  - ✓ 一类是基于边缘的形状表示
  - ✓ 另一类是基于区域的形状表示
- 描述方法
  - ✓ 傅里叶描述子:利用傅里叶变换的边界作为形状特征
  - ✓ 不变矩描述子:利用基于区域的矩作为形状特征
  - ✓ 有限元素法、旋转函数、小波描述子等

基于内容的图像检索



- 基于形状的图像检索的应用
  - ✓ 商标图像检索
  - ✓ 指纹图像检索
  - ✓ 建筑图纸检索
  - ✓ 机械图纸检索
  - ✓ 医学图像检索
  - ✓ ……



商标检索

Query Image		Relevant Trademarks Retrieved in Top 20
(1)		
(2)		
(3)		
(4)		
(5)		
(6)		
(7)		
(8)		
(9)		
(10)		

基于内容的图像检索



- 颜色特征总结

  - ✓ 是一种全局特征，描述了图像或图像区域所对应景物的表面性质
  - ✓ 由于颜色对图像或图像区域的方向、大小等变化不敏感，所以颜色特征不能很好地捕捉图像中对象的局部特征
  - ✓ 颜色直方图表示的优缺点
    - ➢ 优点：不受图像旋转和平移变化的影响，归一化可不受图像尺度变化的影响
    - ➢ 缺点：没有颜色空间分布的信息？有没有办法？

基于内容的图像检索



- 纹理特征总结

  ✓ 也是一种全局特征，描述了图像或图像区域所对应景物的表面性质

  ✓ 仅仅利用纹理特征是无法获得高层次图像内容

  ✓ 不像颜色特征基于像素点，纹理需要在包含多个像素点的区域中进行统计计算

  ✓ 具有旋转不变性，并且对于噪声有较强的抵抗能力

基于内容的图像检索



- 纹理特征总结
  - ✓缺点
    - ➢当图像的分辨率发生变化，所计算出来的纹理可能会有较大偏差
    - ➢可能受到光照、反射情况的影响，用于检索时，这些虚假的纹理会造成误导
  - ✓ 可用于检索具有粗细、疏密等方面较大差别的纹理图像
  - ✓ 如果纹理之间的粗细、疏密相差不大，则难于应用

基于内容的图像检索



- 形状特征总结
  - ✓ 可以有效利用图像中感兴趣的目标来进行检索
  - ✓ 存在的问题：
    - ➢ 缺乏比较完善的数学模型
    - ➢ 如果目标有变形时，检索结果不太可靠
    - ➢ 仅仅描述了目标局部的性质，要全面描述目标常常对计算时间和存储量有较高的要求
    - ➢ 许多形状特征所反映的目标形状信息与人的直观感觉不完全一致

不同特征的相对特点



- 颜色vs纹理
  - ✓ 颜色特征充分利用了图像的色彩信息，而纹理特征只利用了图像的灰度信息（彩色纹理特征不多见）
  - ✓ 颜色特征侧重于图像整体信息的描述，而纹理特征更偏重于局部
- 颜色vs形状
  - ✓ 颜色特征多具有平移、旋转和尺度不变性，而不少形状特征（如边缘方向）只具有平移不变性

不同特征的相对特点



- 纹理vs形状
  - ✓ 通常纹理特征比较容易获得，而形状特征的计算常比较复杂
- 颜色vs形状
  - ✓ 颜色特征多具有平移、旋转和尺度不变性，而不少形状特征（如边缘方向）只具有平移不变性

基于内容的图像检索



- 为什么需要基于内容的图像检索?
- 查询方式，查询demo，现有系统简介
- 具体内容
  - ✓ 特征提取
  - ✓ 相似度匹配
  - ✓ 相关反馈
  - ✓ 索引结构
- MPEG-7介绍：性能评价等
- 思考的几个问题？

相关反馈



- 为什么需要反馈
  - ✓ 目前所提取的描述特征还不能完整或全面地描述图像的内容
    - ➢ 很难说不同特征的不同表达，哪种最好
    - ➢ 对高层概念和视觉感知都很难建立有效模型
  - ✓ 人类不具有关于底层特征的全面知识
  - ✓ 所以，仅通过一次检索就找到需要的图像是不现实的
  - ✓ 将用户的特殊要求反馈给系统，以使检索更有效且更接近用户的要求

相关反馈



- 为什么需要反馈
  - ✓ 基于底层视觉特征的检索系统带来的主要问题
    - ➢ 人类高层语义特征和计算机底层特征表示的矛盾
      - 十五的月亮，容易想到颜色、亮度和形状
      - 漂亮的衣服，很难想到颜色、长短或尺寸
    - ➢ 人类视觉感知的主观性——因人而异
      - 有人关注图像的区域色彩而有人关注区域布局
      - 可能有人注重色调而有人注重饱和度

相关反馈



- 相关反馈的特点
  - ✓ 最先用于文字检索，已有30多年的历史[Rocchio 1971]
  - ✓ 利用用户先前的检索结果信息进行反馈来自动调节当前查询，也可借助人机交互细化用底层特征表达的高层查询
  - ✓ 使用相关反馈来建立高层概念和底层特征之间的联系

相关反馈



- 相关反馈的过程
  - ✓ 用户进行一次初步查询，系统返回检索结果
  - ✓ 用户判断结果，标记满意的结果为正例，标记不满意的结果为反例
  - ✓ 系统根据用户标记结果进行自我调整（如调整搜索策略、算法参数、不同特征的权重等），进行新一轮检索
  - ✓ 根据用户要求，反复进行

α

相关反馈



- 相关反馈的研究
  - ✓ 查询点移动：改进对理想查询点的估计，具体是将该查询点移向好的范例点而远离坏的范例点

Rocchio提出的如下公式

$$Q_{i+1} = \alpha Q_i + \beta \sum_{rel} \frac{D_i}{|D_i|} - \gamma \sum_{nonrel} \frac{D_i}{|D_i|}$$

其中 Q_i和Q_{i+1}分别对应查询迭代中第i次和第i+1次的查询点位置；D_i是特征向量，$|D_i|$是 D_i 的基数，rel和$nonrel$表示相关文档和非相关文档，

α, β 和 γ 是适当的参数，已在MARS中被实现

α

相关反馈



- 相关反馈的研究

  ✓ 再加权/调整权重：提高那些有助于检索相关图像的特征的权重，同时降低有碍于检索相关图像的特征的权重。MARS系统实现了被称为标准背离方法的再加权方法的较小的改进

基于内容的图像检索



- 为什么需要基于内容的图像检索?
- 查询方式，查询demo，现有系统简介
- 具体内容
    - ✓ 特征提取
    - ✓ 相似度匹配
    - ✓ 相关反馈
    - ✓ 索引结构
- MPEG-7介绍：性能评价等
- 思考的几个问题?

基于内容的图像检索



- 为什么需要索引结构?
  - ✓ 海量的图像内容
  - ✓ 图像检索的速度要求
- 可用的索引结构
  - ✓ 满足距离定义的索引结构，如使用欧式距离度量图像的相似度
  - ✓ 距离的定义

基于内容的图像检索



- 可用的索引结构
  - ✓ 距离的定义（续）：

    ① $d(x, y) = d(y, x)$

    ② $d(x, x) = 0$

    ③ $0 < d(x, y) < \infty, \quad x \neq y$

    ④ $d(x, y) \leqslant d(x, z) + d(z, y)$ （三角不等式）

  - ✓ 图像相似度的实际情况
  - ✓ 挑战性问题

    如何兼顾检索正确性和速度的要求，研究图像检索的索引结构（包括视频镜头和视频片断的索引结构）

基于内容的图像检索



- 为什么需要基于内容的图像检索?
- 查询方式，查询demo，现有系统简介
- 具体内容
  - ✓ 特征提取
  - ✓ 相似度匹配
  - ✓ 相关反馈
  - ✓ 索引结构
- MPEG-7介绍：性能评价等
- 思考的几个问题?

基于内容的图像检索



- MPEG-7
  - ✓ Multimedia Content Description Interface, 多媒体内容描述接口
  - ✓ 包括静止图像、运动图像、计算机图形、3-D模型、动画、语言、声音等
  - ✓ 建立对不同多媒体信息（主要为音频视频）的标准描述（包括指定一组描述符和描述方案）
  - ✓ 仅仅是描述有关内容的信息，但并不是内容本身

基于内容的图像检索



- MPEG-7标准的4个特点
  - ✓ 通用：能够描述各种应用环境的多媒体内容
  - ✓ 基于目标的数据模型：能够分别描述单个目标，如场景和场景中的目标
  - ✓ 集成：将底层的特征和高层的描述集成到同一个结构中，结合了它们的描述能力
  - ✓ 可扩展：描述定义语言保证了可将MPEG-7标准扩展到新的应用领域，满足新的需求并集成新的描述工具

基于内容的图像检索



分析　　MPEG-7　　提取

数据 → 特征提取 → 内容描述 → 搜索引擎 → 信息

数据库

基于内容的图像检索



- MPEG-7不包括特征提取的工具和算法，也不包括搜索引擎
- 让产业界开展竞争，以得出最好的结果
- MPEG-7仅仅标准化内容描述工具，而将内容分析工具（如特征提取）和内容使用工具（如搜索引擎和内容过滤）交给产业界
- 仅仅将对多媒体描述的格式（包括句法和语义）归于需要规范的部分，而将其它归于非规范的部分

MPEG-7对试验的性能评价



- ANMRR（Average Normalized Modified Retrieval Rank），平均归一化调整后的检索秩
- AR（Average Recall）
- $AR, ANMRR \in [0,1]$
- AR值越高，查全性能越好
- ANMRR值越低，查准性能越好，说明更多正确的结果排在前面

MPEG-7对试验的性能评价



- 首先挑选出一个查询图像的集合Q，对每个查询图像主观地选取一组视觉相似的图像作为标准，正确答案(Ground Truth)
- 设查询图像q的相似图像的个数为$ng(q)$。对于查询图像q，检索结果的截断值K定义为

$\min\{4\times ng(q), 2\times GTM\}$，其中，$GTM$是在所有查询图像中最大的相似图像个数，即$GTM=max\{ng(q)\}$

- 对于查询图像q，在前K个检索结果中正确的个数记为$nr(q)$，漏掉的个数记为$M(q)=ng(q)-nr(q)$。查全率记为$R(q)=nr(q)/ng(q)$

MPEG-7对试验的性能评价



- 每个正确答案在检索结果中都有一个秩 *(rank)* $r(i)$，$i=1,\ldots,ng(q)$。在前K个检索结果中，正确的图像的秩$r(i)$就是它的序号，其余被漏掉的图像的秩$r(i)$都设定为K+1
- 对于某个查询图像q，它的平均检索秩和调整后的检索秩分别定义为

$$ARR(q) = \sum_{i=1}^{ng(q)} \frac{r(i)}{ng(q)}$$

$$MRR(q) = ARR(q) - \frac{ng(q)}{2} - 0.5$$

MPEG-7对试验的性能评价



- 将$MRR(q)$归一化至[0, 1]范围内，得到归一化调整后的检索秩$NMRR(q)$：

$$NMRR(q) = \frac{MRR(q)}{K - \frac{ng(q)}{2} + 0.5}$$

- 对Q中所有的查询图像q的$NMRR(q)$和$R(q)$作平均，得到$ANMRR$和AR：

$$ANMRR = \frac{1}{Q}\sum_{q=1}^{Q} NMRR(q)$$

$$AR = \frac{1}{Q}\sum_{q=1}^{Q} R(q)$$

MPEG-7对试验的性能评价



- AR与ANMRR说明
  - ✓ 不仅适用于评价图像，也适用于镜头、片断、文本检索
  - ✓ 不能计算“哪个更像”，只能统计正确结果的顺序

基于内容的图像检索



- 为什么需要基于内容的图像检索?
- 查询方式，查询demo，现有系统简介
- 具体内容
  - ✓ 特征提取
  - ✓ 相似度匹配
  - ✓ 相关反馈
- MPEG-7介绍：性能评价等
- 思考的几个问题?



思考的几个问题

1. 基于查询例子的检索是好的吗？查询图像从何而来？如何用于搜索引擎？

思考的几个问题

花？



2. 如何解决人类高层语义和计算机底层特征表示的矛盾？如何实现语义检索？中间层？

思考的几个问题



3. 个人照片如何管理和索引？数码相机，…
   - ✓ 地点：在西藏的？在香港的？
   - ✓ 人物：父母的？女朋友的？
   - ✓ 时间：2004年五一的？小时候的照片？
   - ✓ 景色：泰山？黄果树瀑布？
   - ✓ ……






思考的几个问题



主要参考文献：

- 章毓晋，基于内容的视觉信息检索，科学出版社，2003年。



任何问题？

研究生课程



数字图像处理

Digital Image Processing

彭宇新

北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

傅里叶变换



- 傅里叶变换
  - ✓ 傅里叶变换及其反变换
  - ✓ 傅里叶变换的性质
  - ✓ 快速傅里叶变换（FFT）

傅里叶变换



- 为什么要在频率域研究图像增强

  ✓ 可以利用频率成分和图像外表之间的对应关系。一些在空间域表述困难的增强任务，在频率域中变得非常普通

  ✓ 滤波在频率域更为直观，它可以解释空间域滤波的某些性质

  ✓ 可以在频率域指定滤波器，做反变换，然后在空间域使用结果滤波器作为空间域滤波器的指导

  ✓一旦通过频率域试验选择了空间滤波，通常实施都在空间域进行

傅里叶变换



- 一维连续傅里叶变换及反变换
  - ✓ 单变量连续函数f(x)的傅里叶变换F(u)定义为

$$F(u)=\int_{-\infty}^{\infty} f(x)e^{-j2\pi ux}dx$$

其中，$j=\sqrt{-1}$

  - ✓ 给定F(u)，通过傅里叶反变换可以得到f(x)

$$f(x)=\int_{-\infty}^{\infty} F(u)e^{j2\pi ux}du$$

傅里叶变换



- 二维连续傅里叶变换及反变换
  - ✓ 二维连续函数f(x, y)的傅里叶变换F(u, v)定义为

$$F(u,v)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-j2\pi(ux+vy)}dxdy$$

  - ✓ 给定F(u, v), 通过傅里叶反变换可以得到f(x, y)

$$f(x,y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)e^{j2\pi(ux+vy)}dudv$$

傅里叶变换



- 一维离散傅里叶变换(DFT)及反变换
  - ✓ 单变量离散函数f(x)(x=0,1,2,...,M-1)的傅里叶变换F(u)定义为

$$F(u)=\frac{1}{M}\sum_{x=0}^{M-1}f(x)e^{-j2\pi ux/M}$$

u=0, 1, 2, ..., M-1

  - ✓ 给定F(u),通过傅里叶反变换可以得到f(x)

$$f(x)=\sum_{u=0}^{M-1}F(u)e^{j2\pi ux/M}$$

x=0, 1, 2, ..., M-1

傅里叶变换



- 一维离散傅里叶变换及反变换
  - ✓ 从欧拉公式 $e^{j\theta} = \cos\theta + j\sin\theta$

$$F(u) = \frac{1}{M}\sum_{x=0}^{M-1} f(x) e^{j(-2\pi ux)/M}$$

$$= \frac{1}{M}\sum_{x=0}^{M-1} f(x)\left(\cos(-2\pi ux)/M + j\sin(-2\pi ux)/M\right)$$

$$= \frac{1}{M}\sum_{x=0}^{M-1} f(x)\left(\cos 2\pi ux/M - j\sin 2\pi ux/M\right)$$

傅里叶变换



- 傅里叶变换的极坐标表示

$$F(u)=|F(u)|e^{-j\phi(u)}$$

✓ 幅度或频率谱为

$$|F(u)|=\left[R(u)^2+I(u)^2\right]^{\frac{1}{2}}$$

R(u)和I(u)分别是F(u)的实部和虚部

✓ 相角或相位谱为

$$\phi(u)=\arctan\left[\frac{I(u)}{R(u)}\right]$$

傅里叶变换



- 傅里叶变换的极坐标表示
  - 功率谱为

$$P(u)=|F(u)|^2=R(u)^2+I(u)^2$$

- f (x) 的离散表示

$$f(x)\cong f(x_0+x\Delta x) \qquad x=0,1,2,\ldots,M-1$$

- F (u) 的离散表示

$$F(u)\cong F(u\Delta u) \qquad u=0,1,2,\ldots,M-1$$

傅里叶变换



- 二维离散傅里叶变换及反变换
  - ✓ 图像尺寸为M×N的函数f(x, y)的DFT为

$$F(u,v)=\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)e^{-j2\pi(ux/M+vy/N)}$$

u=0, 1, 2, …, M-1,　　v=0, 1, 2, …, N-1

  - ✓ 给出F(u, v), 可通过反DFT得到f(x, y),

$$f(x,y)=\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}F(u,v)e^{j2\pi(ux/M+vy/N)}$$

x=0, 1, 2, …, M-1,　　y=0, 1, 2, …, N-1

注：u和v是频率变量，x和y是空间或图像变量

傅里叶变换



- 二维DFT的极坐标表示

$$F(u,v)=|F(u,v)|e^{-j\phi(u,v)}$$

✓ 幅度或频率谱为

$$|F(u,v)|=\left[R(u,v)^2+I(u,v)^2\right]^{\frac{1}{2}}$$

R(u, v)和I(u, v)分别是F(u, v)的实部和虚部

✓ 相角或相位谱为

$$\phi(u,v)=\arctan\left[\frac{I(u,v)}{R(u,v)}\right]$$

傅里叶变换



- 二维DFT的极坐标表示
  - ✓ 功率谱为

$$P(u,v)=\left|F(u,v)\right|^2=R(u,v)^2+I(u,v)^2$$

- F(u, v)的原点变换

$$\Im\left[f(x,y)(-1)^{x+y}\right]=F(u-M/2,v-N/2)$$

  - ✓ 用$(-1)^{x+y}$乘以f(x, y)，将F(u, v)原点变换到频率坐标下的(M/2，N/2)，它是M×N区域的中心
  - ✓ u=0, 1, 2, …, M-1,　　　v=0, 1, 2, …, N-1

傅里叶变换



- F(0, 0)表示

$$F(0,0)=\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)$$

这说明：假设f(x, y)是一幅图像，在原点的傅里叶变换等于图像的平均灰度级

傅里叶变换



- 如果f(x, y)是实函数，它的傅里叶变换是对称的，即

$$F(u,v)=F(-u,-v)$$

- 傅里叶变换的频率谱是对称的

$$|F(u,v)|=|F(-u,-v)|$$

傅里叶变换



- 傅里叶变换
  - ✓ 傅里叶变换及其反变换
  - ✓ 傅里叶变换的性质
  - ✓ 快速傅里叶变换（FFT）

傅里叶变换



- 二维傅里叶变换的性质

1. 平移性质
2. 分配律
3. 尺度变换（缩放）
4. 旋转性
5. 周期性和共轭对称性
6. 平均值
7. 可分性
8. 卷积
9. 相关性

傅里叶变换



1. 傅里叶变换对的平移性质

以⇔表示函数和其傅里叶变换的对应性

$$f(x,y)e^{j2\pi(u_0x/M+v_0y/N)} \Leftrightarrow F(u-u_0,v-v_0) \quad (1)$$

$$f(x-x_0,y-y_0) \Leftrightarrow F(u,v)e^{-j2\pi(ux_0/M+vy_0/N)} \quad (2)$$

- ✓ 公式（1）表明将f(x, y)与一个指数项相乘就相当于把其变换后的频域中心移动到新的位置
- ✓ 公式（2）表明将F(u, v)与一个指数项相乘就相当于把其变换后的空域中心移动到新的位置
- ✓ 公式（2）表明对f(x, y)的平移不影响其傅里叶变换的幅值

傅里叶变换



1. 傅里叶变换对的平移性质（续）

当u_0=M/2且v_0=N/2,

$$e^{j2\pi(u_0x/M+v_0y/N)}=e^{j\pi(x+y)}=(-1)^{x+y}$$

带入（1）和（2），得到

$$f(x,y)(-1)^{x+y}\Leftrightarrow F(u-M/2,v-N/2)$$

$$f(x-M/2,y-N/2)\Leftrightarrow F(u,v)(-1)^{u+v}$$

傅里叶变换



2. 分配律

根据傅里叶变换的定义，可以得到

$$\Im[f_1(x,y)+f_2(x,y)]=\Im[f_1(x,y)]+\Im[f_2(x,y)]$$

$$\Im[f_1(x,y)\bullet f_2(x,y)]\neq\Im[f_1(x,y)]\bullet\Im[f_2(x,y)]$$

上述公式表明：傅里叶变换对加法满足分配律，但对乘法则不满足

傅里叶变换



3\. 尺度变换（缩放）

给定2个标量a和b，可以证明对傅里叶变换下列2个公式成立

$$af(x,y) \Leftrightarrow aF(u,v)$$

$$f(ax,by) \Leftrightarrow \frac{1}{|ab|}F(u/a,v/b)$$

傅里叶变换



4. 旋转性

引入极坐标 $x=r\cos\theta, y=r\sin\theta, u=\omega\cos\varphi, v=\omega\sin\varphi$

将f (x, y) 和F (u, v) 转换为 $f(r,\theta)$ 和 $F(\omega,\varphi)$。将它们带入傅里叶变换对得到

$$f(r,\theta+\theta_0)\Leftrightarrow F(\omega,\varphi+\theta_0)$$

- ✓ f (x, y) 旋转角度 θ_0，F (u, v) 也将转过相同的角度
- ✓ F (u, v) 旋转角度 θ_0，f (x, y) 也将转过相同的角度

傅里叶变换



5. 周期性和共轭对称性

$$F(u,v) = F(u+M,v) = F(u,v+N) = F(u+M,v+N)$$

$$f(x,y) = f(x+M,y) = f(x,y+N) = f(x+M,y+N)$$

上述公式表明

- ✓ 尽管F(u, v)对无穷多个u和v的值重复出现，但只需根据在任一个周期里的N个值就可以从F(u, v)得到f(x, y)
- ✓ 只需一个周期里的变换就可将F(u, v)在频域里完全确定
- ✓ 同样的结论对f(x, y)在空域也成立

傅里叶变换



5. 周期性和共轭对称性

如果f(x, y)是实函数，则它的傅里叶变换具有共轭对称性

$$F(u,v)=F^*(-u,-v)$$

$$|F(u,v)|=|F(-u,-v)|$$

其中，F*(u, v)为F(u, v)的复共轭。

- 复习：当两个复数实部相等,虚部互为相反数时,这两个复数叫做互为共轭复数.



周期性和共轭对称性举例

- 对于一维变换F(u)，周期性是指F(u)的周期长度为M，对称性是指频谱关于原点对称

半周期的傅里叶频谱　　全周期的傅里叶频谱

一幅二维图像的傅里叶频谱　　中心化的傅里叶频谱

FIGURE 4.34 (a) Fourier spectrum showing back-to-back half periods in the interval $[0, M-1]$. (b) Shifted spectrum showing a full period in the same interval. (c) Fourier spectrum of an image, showing the same back-to-back properties as (a), but in two dimensions. (d) Centered Fourier spectrum.

傅里叶变换



6. 分离性

$$F(u,v)=\frac{1}{M}\sum_{x=0}^{M-1}e^{-j2\pi ux/M}\frac{1}{N}\sum_{y=0}^{N-1}f(x,y)e^{-j2\pi vy/N}$$

$$=\frac{1}{M}\sum_{x=0}^{M-1}e^{-j2\pi ux/M}F(x,v)$$

F(x, v)是沿着f(x, y)的一行所进行的傅里叶变换。当x=0, 1, …, M-1，沿着f(x, y)的所有行计算傅里叶变换。

傅里叶变换



6. 分离性——二维傅里叶变换的全过程

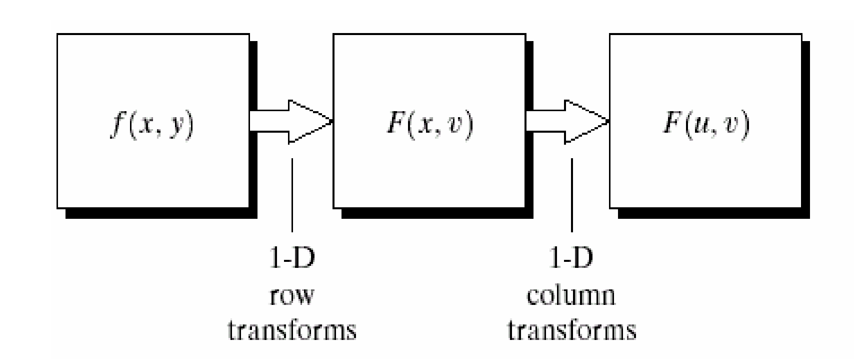


FIGURE 4.35 Computation of the 2-D Fourier transform as a series of 1-D transforms.

- ✓ 先通过沿输入图像的每一行计算一维变换
- ✓ 再沿中间结果的每一列计算一维变换
- ✓ 可以改变上述顺序，即先列后行
- ✓ 上述相似的过程也可以计算二维傅里叶反变换

傅里叶变换



7. 平均值

由二维傅里叶变换的定义

$$F(u,v)=\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)e^{-j2\pi(ux/M+vy/N)}$$

所以 $F(0,0)=\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)$

而 $\bar{f}(x,y)=\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)$

傅里叶变换



7. 平均值

所以

$$\bar{f}(x,y) = F(0,0)$$

上式说明：如果f(x, y)是一幅图像，在原点的傅里叶变换即等于图像的平均灰度级

傅里叶变换



8. 卷积理论

大小为M×N的两个函数f(x, y)和h(x, y)的离散卷积

$$f(x,y)*h(x,y)=\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}f(m,n)h(x-m,y-n)$$

卷积定理

$$f(x,y)*h(x,y)\Leftrightarrow F(u,v)H(u,v)$$

$$f(x,y)h(x,y)\Leftrightarrow F(u,v)*H(u,v)$$

傅里叶变换



9. 相关性理论

大小为M×N的两个函数f(x, y)和h(x, y)的相关性定义为

$$f(x,y)\circ h(x,y)=\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}f^{*}(m,n)h(x+m,y+n)$$

f*表示f的复共轭。对于实函数， f*＝f

相关定理

$$f(x,y)\circ h(x,y)\Leftrightarrow F^{*}(u,v)H(u,v)$$

$$f^{*}(x,y)h(x,y)\Leftrightarrow F(u,v)\circ H(u,v)$$

傅里叶变换



- 自相关理论

$$f(x,y)\circ f(x,y) \Leftrightarrow |F(u,v)|^2 = R(u,v)^2 + I(u,v)^2$$

$$|f(x,y)|^2 \Leftrightarrow F(u,v)\circ F(u,v)$$

注：复数和它的复共轭的乘积是复数模的平方

傅里叶变换



- 卷积和相关性理论总结
  - ✓ 卷积是空间域过滤和频率域过滤之间的纽带
  - ✓ 相关的重要应用在于匹配：确定是否有感兴趣的物体区域
    - ➢ f(x, y)是原始图像
    - ➢ h(x, y)作为感兴趣的物体或区域（模板）
    - ➢ 如果匹配，两个函数的相关值会在h找到f中相应点的位置上达到最大

相关性匹配举例



图像f(x, y)

模板h(x, y)

延拓图像f(x, y)

延拓图像h(x, y)

相关函数图像

通过相关图像最大值的水平灰度剖面图

傅里叶变换



- 傅里叶变换
  - ✓ 傅里叶变换及其反变换
  - ✓ 傅里叶变换的性质
  - ✓ 快速傅里叶变换(FFT)
    - ➢ 只考虑一维的情况，根据傅里叶变换的分离性可知，二维傅里叶变换可由连续2次一维傅里叶变换得到

快速傅里叶变换(FFT)



- 为什么需要快速傅里叶变换？

$$F(u)=\frac{1}{M}\sum_{x=0}^{M-1}f(x)e^{-j2\pi ux/M} \qquad u=0,1,2,\ldots,M-1$$

✓ 对u的M个值中的每一个都需进行M次复数乘法(将f(x)与 $e^{-j2\pi ux/M}$ 相乘)和M-1次加法，即复数乘法和加法的次数都正比于M^2

✓ 快速傅里叶变换(FFT)则只需要$M\log_2 M$次运算

✓ FFT算法与原始变换算法的计算量之比是$\log_2 M/M$，如$M=1024\approx10^3$,则原始变换算法需要10^6次计算，而FFT需要10^4次计算，FFT与原始变换算法之比是1：100

快速傅里叶变换(FFT)



- FFT算法基本思想

  FFT算法基于一个叫做逐次加倍的方法。通过推导将原始傅里叶转换成两个递推公式

$$F(u)=\frac{1}{M}\sum_{x=0}^{M-1} f(x)e^{-j2\pi ux/M} \quad u=0,1,2,\ldots,M-1$$

$$F(u)=\frac{1}{2}\left[F_{even}(u)+F_{odd}(u)W_{2k}^{u}\right]$$

$$F(u+K)=\frac{1}{2}\left[F_{even}(u)-F_{odd}(u)W_{2k}^{u}\right]$$

快速傅里叶变换(FFT)



- FFT算法基本思想

$$F(u)=\frac{1}{2}\left[F_{even}(u)+F_{odd}(u)W_{2k}^{u}\right] \qquad u=0,1,2,\dots,M-1$$

$$F(u+K)=\frac{1}{2}\left[F_{even}(u)-F_{odd}(u)W_{2k}^{u}\right]$$

其中：M = 2K

$F_{even}(u)$、$F_{odd}(u)$是K个点的傅里叶值

快速傅里叶变换(FFT)



- FFT公式推导

  FFT算法基于一个叫做逐次加倍的方法。为方便起见用下式表达离散傅立叶变换公式

$$F(u)=\frac{1}{M}\sum_{x=0}^{M-1}f(x)e^{-j2\pi ux/M}$$

$$=\frac{1}{M}\sum_{x=0}^{M-1}f(x)W_M^{ux}$$

这里 $W_M=e^{-j2\pi/M}$ 是一个常数

快速傅里叶变换(FFT)



假设M的形式是

$$M = 2^n$$

n为正整数。因此，M可以表示为

$$M = 2K$$

将M=2K带入上式

$$F(u) = \frac{1}{2K}\sum_{x=0}^{2K-1} f(x) W_{2K}^{ux}$$

$$= \frac{1}{2}\left[\frac{1}{K}\sum_{x=0}^{K-1} f(2x) W_{2K}^{u(2x)} + \frac{1}{K}\sum_{x=0}^{K-1} f(2x+1) W_{2K}^{u(2x+1)}\right]$$

快速傅里叶变换(FFT)



推导：因为

$$W_M = e^{-j2\pi/M}$$

所以

$$W_{2K}^{2ux} = e^{-j2\pi(2ux)/2K} = e^{-j2\pi(ux)/K} = W_K^{ux}$$

带入上式有

$$F(u) = \frac{1}{2}\left[\frac{1}{K}\sum_{x=0}^{K-1} f(2x)W_K^{ux} + \frac{1}{K}\sum_{x=0}^{K-1} f(2x+1)W_K^{ux}W_{2K}^{u}\right]$$



快速傅里叶变换(FFT)



定义两个符号

$$F_{even}(u) = \frac{1}{K}\sum_{x=0}^{K-1} f(2x) W_K^{ux} \qquad u = 0,1,2,\ldots,K-1$$

$$F_{odd}(u) = \frac{1}{K}\sum_{x=0}^{K-1} f(2x+1) W_K^{ux}$$

快速傅里叶变换(FFT)



得到FFT的第一个公式

$$F(u)=\frac{1}{2}\left[F_{even}(u)+F_{odd}(u)W_{2K}^{u}\right]$$

该公式说明F(u)可以通过奇部和偶部之和来计算

快速傅里叶变换(FFT)



推导：

$$W_K^{u+K} = e^{-j2\pi(u+K)/K}$$

$$= e^{-j2\pi u/K} e^{-j2\pi}$$

$$= W_K^u e^{j\pi(-2)} = W_K^u (-1)^{(-2)} = W_K^u$$

$$W_{2K}^{u+K} = e^{-j2\pi(u+K)/2K}$$

$$= e^{-j2\pi u/2K} e^{-j\pi}$$

$$= W_{2K}^u e^{j\pi(-1)} = W_{2K}^u (-1)^{(-1)} = -W_{2K}^u$$

快速傅里叶变换(FFT)



$$F(u+K)=\frac{1}{2K}\sum_{x=0}^{2K-1}f(x)W_{2K}^{(u+K)x}$$

$$=\frac{1}{2}\left[\frac{1}{K}\sum_{x=0}^{K-1}f(2x)W_{2K}^{(u+K)(2x)}+\frac{1}{K}\sum_{x=0}^{K-1}f(2x+1)W_{2K}^{(u+K)(2x+1)}\right]$$

$$=\frac{1}{2}\left[\frac{1}{K}\sum_{x=0}^{K-1}f(2x)W_{K}^{(u+K)x}+\frac{1}{K}\sum_{x=0}^{K-1}f(2x+1)W_{K}^{(u+K)x}W_{2K}^{(u+K)}\right]$$

$$=\frac{1}{2}\left[\frac{1}{K}\sum_{x=0}^{K-1}f(2x)W_{K}^{ux}+\frac{1}{K}\sum_{x=0}^{K-1}f(2x+1)W_{K}^{ux}\left(-W_{2K}^{u}\right)\right]$$

$$=\frac{1}{2}\left[F_{even}(u)-F_{odd}(u)W_{2K}^{u}\right]$$

快速傅里叶变换(FFT)



得到FFT的第二个公式

$$F(u+K)=\frac{1}{2}\left[F_{even}(u)-F_{odd}(u)W_{2K}^{u}\right]$$

该公式说明F(u＋K)可以通过奇部和偶部之差来计算



快速傅里叶变换(FFT)



- 最后得到FFT的二个公式

$$F(u)=\frac{1}{2}\left[F_{even}(u)+F_{odd}(u)W_{2K}^{u}\right]$$

$$F(u+K)=\frac{1}{2}\left[F_{even}(u)-F_{odd}(u)W_{2K}^{u}\right]$$

快速傅里叶变换(FFT)



- 分析这些表达式得到如下一些有趣的特性：
  - ✓ 一个M个点的变换，能够通过将原始表达式分成两个部分来计算
  - ✓ 通过计算两个（M/2）个点的变换。得 $F_{even}(u)$ 和 $F_{odd}(u)$
  - ✓ 奇部与偶部之和得到F(u)的前(M/2)个值
  - ✓ 奇部与偶部之差得到F(u)的后(M/2)个值。且不需要额外的变换计算

快速傅里叶变换(FFT)



- 归纳快速傅立叶变换的思想：

（1）通过计算两个单点的DFT，来计算两个点的DFT，

（2）通过计算两个双点的DFT，来计算四个点的DFT，...，以此类推

（3）对于任何$N=2^m$的DFT的计算，通过计算两个N/2点的DFT，来计算N个点的DFT

快速傅里叶变换(FFT)



- FFT算法基本思想

FFT算法举例：

设：有函数f(x)，其$N = 2^3 = 8$，有：

$\{f(0), f(1), f(2), f(3), f(4), f(5), f(6), f(7)\}$

计算：

$\{F(0), F(1), F(2), F(3), F(4), F(5), F(6), F(7)\}$

快速傅里叶变换(FFT)



- FFT算法举例

  首先分成奇偶两组：

有：{ f(0)， f(2)， f(4)， f(6) }

{ f(1)， f(3)， f(5)， f(7) }

为了利用递推特性，再分成两组：

有： { f(0)， f(4) }， { f(2)， f(6) }

{ f(1)， f(5) }， { f(3)， f(7) }

快速傅里叶变换(FFT)



- FFT算法实现
  - ✓ 对输入数据的排序可根据一个简单的位对换规则进行
    - ➢如用x表示f(x)的1个自变量值，那么它排序后对应的值可通过把x表示成二进制数并对换各位得到
    - ➢ 例如$N=2^3$, f(6)排序后为f(3), 因为$6=110_2$而$011_2=3$
  - ✓ 把输入数据进行了重新排序，则输出结果是正确的次序。反之不把输入数据进行排序，则输出结果需要重新排序才能得到正确的次序

快速傅里叶变换(FFT)



- FFT算法实现

  地址的排序：——按位倒序规则

  例如：$N = 2^3 = 8$

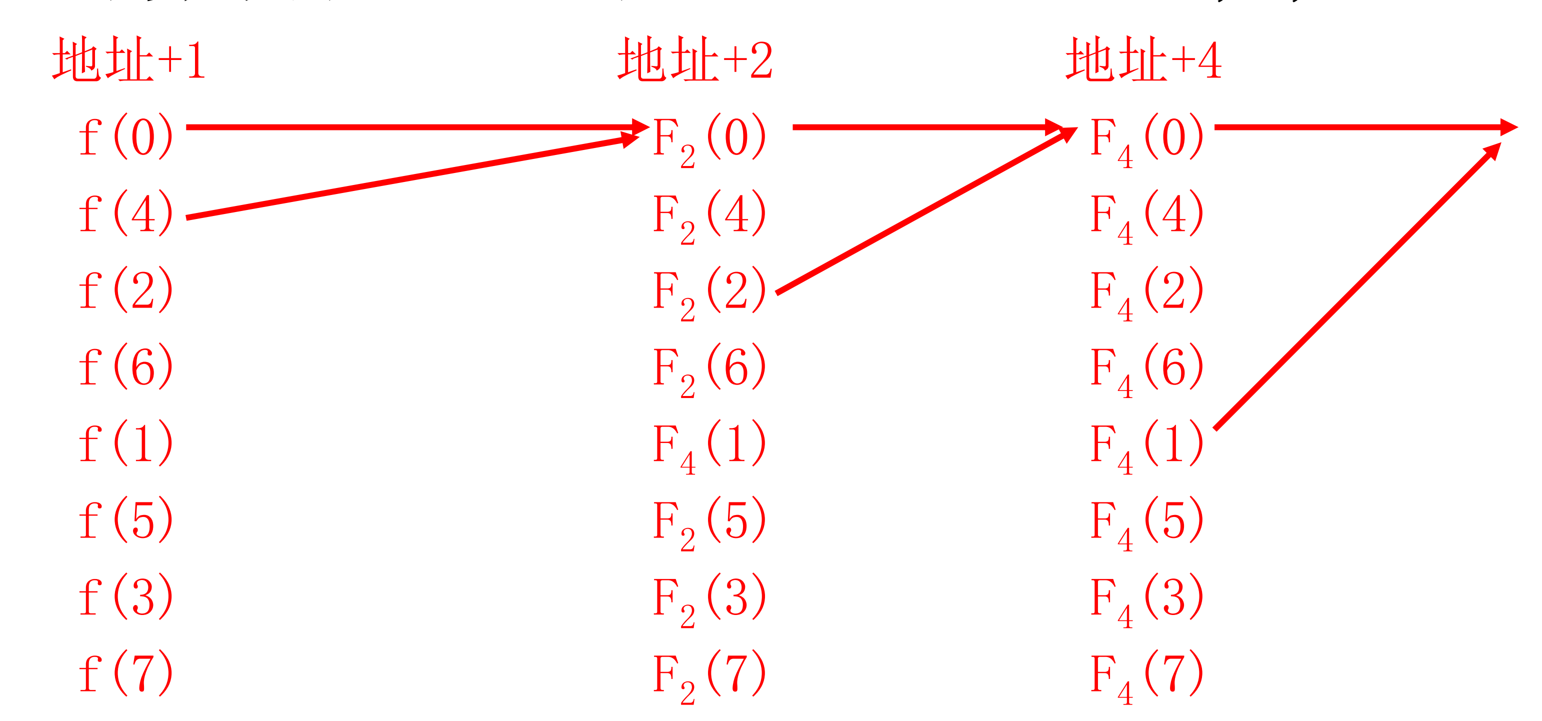
原地址	原顺序	新地址	新顺序
000	f(0)	000	f(0)
001	f(1)	100	f(4)
010	f(2)	010	f(2)
011	f(3)	110	f(6)
100	f(4)	001	f(1)
101	f(5)	101	f(5)
110	f(6)	011	f(3)
111	f(7)	111	f(7)

快速傅里叶变换(FFT)



- FFT算法实现——几个关键点

  2）计算顺序及地址增量：2^n $n = 0, 1, 2...$



数字图像处理（5）



任何问题？

研究生课程



数字图像处理
Digital Image Processing

彭宇新

北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

频率域图像增强



- 频率域滤波
- 频率域平滑（低通）滤波器
- 频率域锐化（高通）滤波器

频率域图像增强



- 为什么要在频率域研究图像增强
  - ✓ 可以利用频率成分和图像外表之间的对应关系。一些在空间域表述困难的增强任务，在频率域中变得非常普通
  - ✓ 滤波在频率域更为直观，它可以解释空间域滤波的某些性质
  - ✓ 给出一个问题，寻找某个滤波器解决该问题，频率域处理对于试验、迅速而全面地控制滤波器参数是一个理想工具
  - ✓ 一旦找到一个特殊应用的滤波器，通常在空间域采用硬件实现它

频率域图像增强



- 傅里叶变换的频率分量和图像空间特征之间的联系
  - ✓ 变化最慢的频率成分(u=v=0)对应一幅图像的平均灰度级

$$F(0,0)=\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)=\bar{f}(x,y)$$

频率域滤波



- 傅里叶变换的频率分量和图像空间特征之间的联系（续）
  - ✓ 当从变换的原点移开时，低频对应着图像的慢变化分量，如图像的平滑部分
  - ✓ 进一步离开原点时，较高的频率对应图像中变化越来越快的灰度级，如边缘或噪声等尖锐部分

频率域滤波



- 频率域的滤波步骤
  1. 用$(-1)^{x+y}$乘以输入图像进行中心变换

$$f(x,y)(-1)^{x+y} \Leftrightarrow F(u-M/2, v-N/2)$$

  2. 计算1中的DFT F(u, v)
  3. 用滤波器函数H(u, v)乘以F(u, v)
  4. 计算3中结果的反DFT
  5. 得到4中结果的实部
  6. 用$(-1)^{x+y}$乘以5中的结果，取消输入图像的乘数

频率域滤波



- 频率域滤波 $G(u,v)=H(u,v)F(u,v)$
  - ✓ H和F的相乘在逐元素的基础上定义，即H的第一个元素乘以F的第一个元素，H的第二个元素乘以F的第二个元素
  - ✓ 一般，F的元素为复数，H的元素为实数
  - ✓ H为零相移滤波器，因为滤波器不改变变换的相位，F中实部和虚部的乘数(H)可以抵消

    相角 $\phi(u,v)=\arctan\left[\dfrac{I(u,v)}{R(u,v)}\right]$

频率域滤波



- 频率域滤波的基本步骤

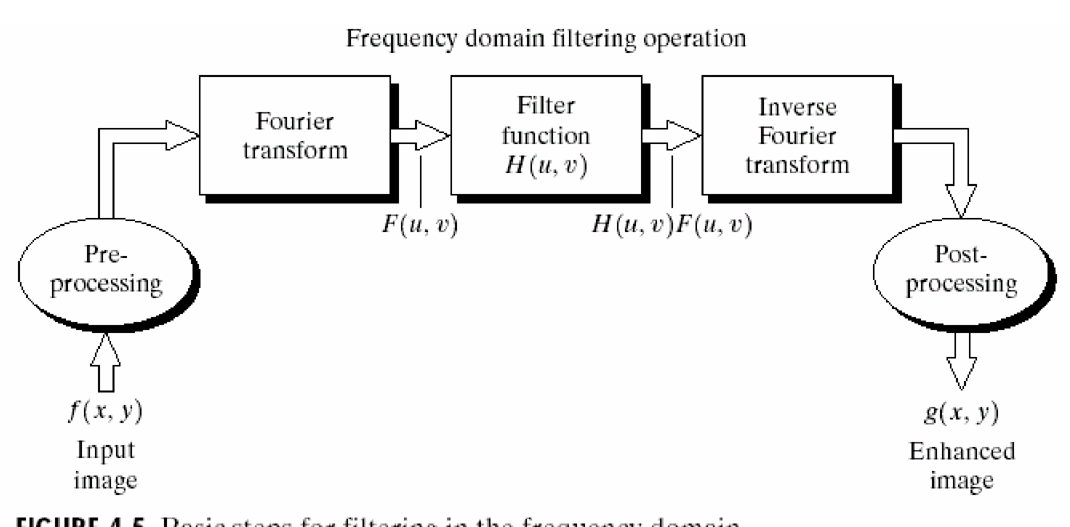


FIGURE 4.5 Basic steps for filtering in the frequency domain.

思想：通过滤波器函数以某种方式来修改图像变换，然后通过取结果的反变换来获得处理后的输出图像

频率域滤波



- 一些基本的滤波器：如何作用于图像？
  - ✓ 陷波滤波器
  - ✓ 低通（平滑）滤波器
  - ✓ 高通（锐化）滤波器

频率域滤波



- 陷波滤波器

$$H(u,v)=\begin{cases}0 & (u,v)=(M/2,N/2)\\ 1 & \text{其它}\end{cases}$$

✓ 设置F(0, 0)=0(结果图像的平均值为零)，而保留其它傅里叶变换的频率成分不变

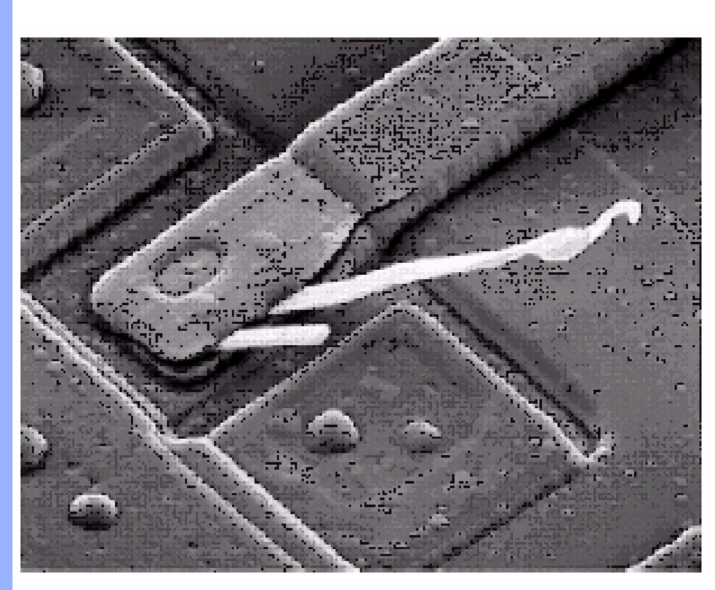
✓ 除了原点处有凹陷外，其它均是常量函数

✓ 由于图像平均值为0而产生整体平均灰度级的降低

✓ 用于识别由特定的、局部化频域成分引起的空间图像效果

陷波滤波器举例









由于图像平均值为0而产生整体平均灰度级的降低

频率域滤波



- 低通滤波器：使低频通过而使高频衰减的滤波器
  - ✓ 被低通滤波的图像比原始图像少尖锐的细节部分而突出平滑过渡部分
  - ✓ 对比空间域滤波的平滑处理，如均值滤波器
- 高通滤波器：使高频通过而使低频衰减的滤波器
  - ✓ 被高通滤波的图像比原始图像少灰度级的平滑过渡而突出边缘等细节部分
  - ✓ 对比空间域的梯度算子、拉普拉斯算子

低通滤波器和高通滤波器举例



原图

低通滤波函数

$H(u,v)$

低通滤波结果：模糊

高通滤波函数

$H(u,v)$

Origin

高通滤波结果：锐化

FIGURE 4.7 (a) A two-dimensional lowpass filter function. (b) Result of lowpass filtering the image in Fig. 4.4(a). (c) A two-dimensional highpass filter function. (d) Result of highpass filtering the image in Fig. 4.4(a).

低通滤波器和高通滤波器举例





- ✓ 因为F(0,0)已被设置为0，所以几乎没有平滑的灰度级细节，且图像较暗
- ✓ 在滤波器中加入常量，以使F(0,0)不被完全消除，如图所示，对滤波器加上一个滤波器高度一半的常数加以改进（高频加强）

频率域滤波



- 空间域滤波和频率域滤波之间的对应关系
  - ✓ 大小为M×N的两个函数f(x, y)和h(x, y)的离散卷积表示为f(x, y)*h(x, y)，定义为

$$f(x,y)*h(x,y)=\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}f(m,n)h(x-m,y-n)$$

  - ✓ 对比空间域滤波：在M×N的图像f上，用m×n的滤波器进行线性滤波

$$g(x,y)=\sum_{s=-a}^{a}\sum_{t=-b}^{b}w(s,t)f(x+s,y+t)$$

频率域滤波



- 卷积定理

$$f(x,y)*h(x,y) \Leftrightarrow F(u,v)H(u,v)$$

  - ✓ 上式说明空间域卷积可以通过F(u, v)H(u, v)的乘积进行反傅里叶变换得到

$$f(x,y)h(x,y) \Leftrightarrow F(u,v)*H(u,v)$$

  - ✓ 说明空间域乘法可以通过频率域的卷积获得
  - ✓ 上述两个公式主要为两个函数逐元素相乘的乘法

频率域滤波



- 定义：在(x_0, y_0)，强度为A的冲激函数表示为

  $A\delta(x-x_0, y-y_0)$，定义为

$$\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} s(x,y)A\delta(x-x_0, y-y_0) = As(x_0, y_0)$$

  - ✓ s(x, y)是冲激函数
  - ✓ 说明： s(x, y)只在(x_0, y_0)处有为A的图像值，其它处的值全为0

频率域滤波



- 下式表明在原点处(0, 0)的单位冲激情况，用 $\delta(x, y)$ 表示

$$\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} s(x, y)\delta(x, y) = s(0,0)$$

- 根据上式，原点处(0, 0)单位冲激的傅里叶变换

$$\Im[\delta(x, y)] = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\delta(x, y)e^{-j2\pi(ux/M+vy/N)}$$

$$= \frac{1}{MN}e^{0} \quad = \boxed{\frac{1}{MN}}$$

频率域滤波



- 假设 $f(x,y)=\delta(x,y)$，根据上式计算原点处 $(0,0)$ 空间域的卷积

$$\delta(x,y)*h(x,y)=\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\delta(m,n)h(x-m,y-n)$$

$$=\frac{1}{MN}h(x-0,y-0)=\boxed{\frac{1}{MN}h(x,y)}$$

- 结论

$$f(x,y)*h(x,y)\Leftrightarrow F(u,v)H(u,v)$$

$$\boxed{\delta(x,y)*h(x,y)\Leftrightarrow \Im[\delta(x,y)]H(u,v)}$$

$$\boxed{h(x,y)\Leftrightarrow H(u,v)}$$

频率域滤波



$$h(x, y) \Leftrightarrow H(u, v)$$

- ✓ 上述公式表明，空间域和频率域中的滤波器组成了傅里叶变换对
- ✓ 给出在频率域的滤波器，可以通过反傅里叶变换得到在空间域对应的滤波器，反之亦然
- ✓ 滤波在频率域中更为直观，但在空间域一般使用更小的滤波器模板
- ✓ 可以在频率域指定滤波器，做反变换，然后在空间域使用结果滤波器作为在空间域构建小滤波器模板的指导

频率域滤波



- 高斯频率域低通滤波器函数

$$H(u) = Ae^{-u^2/2\sigma^2}$$

对应空间域高斯低通滤波器为

$$h(x) = \sqrt{2\pi}\sigma Ae^{-2\pi^2\sigma^2x^2}$$

- 高斯频率域高通滤波器函数

$$H(u) = Ae^{-u^2/2\sigma_1^2} - Be^{-u^2/2\sigma_2^2} \qquad A \geq B, \sigma_1 > \sigma_2$$

对应空间域高斯高通滤波器为

$$h(x) = \sqrt{2\pi}\sigma_1 Ae^{-2\pi^2\sigma_1^2x^2} - \sqrt{2\pi}\sigma_2 Be^{-2\pi^2\sigma_2^2x^2}$$

频率域滤波



频率域高斯低通滤波器　　　　频率域高斯高通滤波器

$\frac{1}{9} \times$

1	1	1
1	1	1
1	1	1

$\frac{1}{16} \times$

1	2	1
2	4	2
1	2	1

−1	−1	−1
−1	8	−1
−1	−1	−1

0	−1	0
−1	4	−1
0	−1	0

a b
c d

FIGURE 4.9 (a) Gaussian frequency domain lowpass filter. (b) Gaussian frequency domain highpass filter. (c) Corresponding lowpass spatial filter. (d) Corresponding highpass spatial filter. The masks shown are used in Chapter 3 for lowpass and highpass filtering.

空间域高斯低通滤波器　　　　空间域高斯高通滤波器

频率域滤波



- 结论（低通滤波器）
  - 当H(u)有很宽的轮廓时(大的σ 值)，h(x)有很窄的轮廓，反之亦然。当σ接近无限时，H(u)趋于常量函数，而h(x)趋于冲激函数
  - 两个低通滤波器的相似之处在于两个域中的值均为正。所以，在空间域使用带正系数的模板可以实现低通滤波
  - 频率域低通滤波器越窄，滤除的低频成分就越多，使得图像就越模糊；在空间域，这意味着低通滤波器就越宽，模板就越大

频率域滤波



- 结论（高通滤波器）
  - ✓ 空间域滤波器有正值和负值，一旦值变为负数，就再也不会变为正数
- 为什么频率域中的内容在空间域要使用小空间模板
  - ✓ 频率域可以凭直观指定滤波器
  - ✓ 空间域滤波效果取决于空间模板的大小

频率域图像增强



- 频率域平滑滤波器
  - ✓ 理想低通滤波器
  - ✓ 巴特沃思低通滤波器
  - ✓ 高斯低通滤波器
  - ✓ 应用实例

频率域图像增强



- 频率域平滑滤波器
  - ✓ 边缘和噪声等尖锐变化处于傅里叶变换的高频部分
  - ✓ 平滑可以通过衰减高频成分的范围来实现
  - ✓ 理想低通滤波器：尖锐
  - ✓ 巴特沃思低通滤波器：处于理想和高斯滤波器之间
  - ✓ 高斯低通滤波器：平滑

频率域图像增强



- 理想低通滤波器

✓ 截断傅里叶变换中的所有高频成分，这些高频成分处于指定距离D_0之外

$$H(u,v)=\begin{cases}1 & D(u,v)\le D_0 \\ 0 & D(u,v)> D_0\end{cases}$$

✓ 频率矩形的中心在(u, v)=(M/2, N/2)，从点(u, v)到中心（原点）的距离如下

$$D(u,v)=\left[\left(u-\frac{M}{2}\right)^2+\left(v-\frac{N}{2}\right)^2\right]^{\frac{1}{2}}$$

频率域图像增强



- 理想低通滤波器

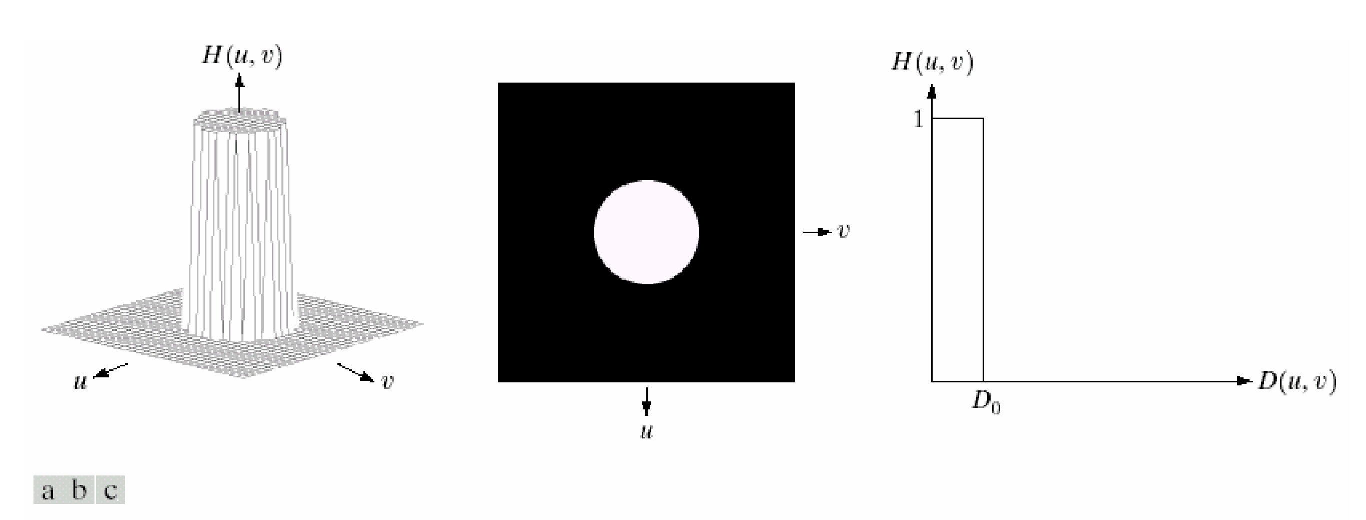


FIGURE 4.10 (a) Perspective plot of an ideal lowpass filter transfer function. (b) Filter displayed as an image. (c) Filter radial cross section.

说明：在半径为D_0的圆内，所有频率没有衰减地通过滤波器，而在此半径的圆之外的所有频率完全被衰减掉

频率域图像增强



- 理想低通滤波器

  ✓ 总图像功率值P_T

$$P_T = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} P(u,v)$$

$$P(u,v) = |F(u,v)|^2 = R(u,v)^2 + I(u,v)^2$$

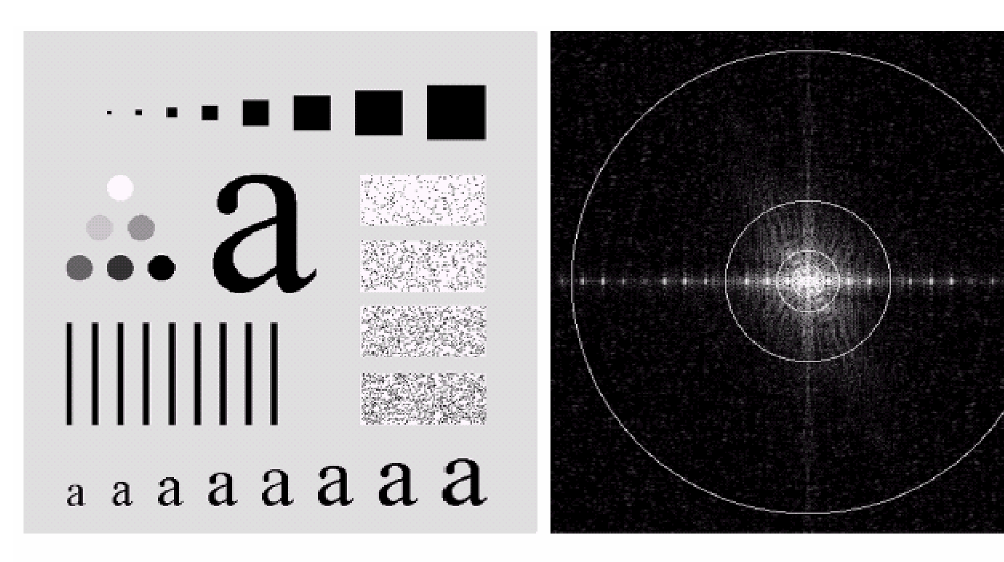
  ✓ 原点在频率矩形的中心，半径为r的圆包含α%的功率，其中

$$\alpha = 100\left[\sum_u \sum_v P(u,v) / P_T\right]$$

理想低通滤波器举例



500×500像素的原图　　　图像的傅里叶频谱



a b

FIGURE 4.11 (a) An image of size 500 × 500 pixels and (b) its Fourier spectrum. The superimposed circles have radii values of 5, 15, 30, 80, and 230, which enclose 92.0, 94.6, 96.4, 98.0, and 99.5% of the image power, respectively.

圆环具有半径5, 15, 30, 80和230个像素

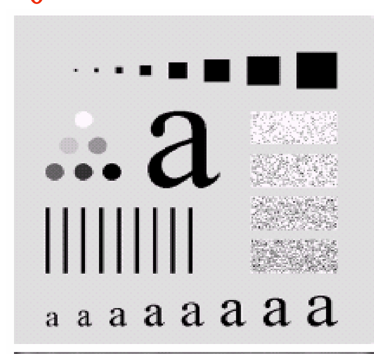
图像功率为92.0%, 94.6%, 96.4%, 98.0%和99.5%

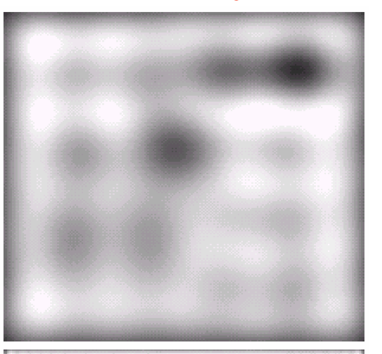
理想低通滤波器举例——具有振铃现象



结论：半径D_0越小，模糊越大；半径D_0越大，模糊越小

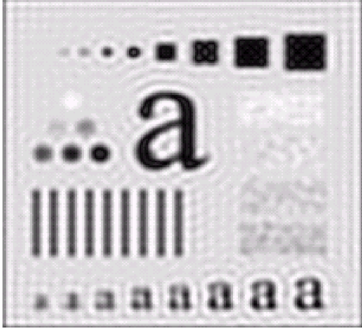
原图





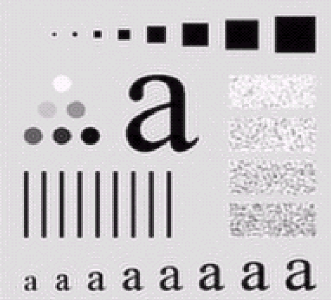
半径是5的理想低通滤波，滤除8%的总功率，模糊说明多数尖锐细节在这8%的功率之内

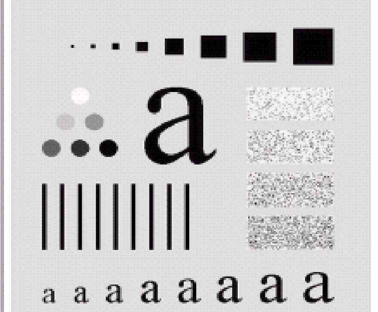
半径是15的理想低通滤波，滤除5.4%的总功率



半径是30的理想低通滤波，滤除3.6%的总功率

半径是80的理想低通滤波，滤除2%的总功率





半径是230的理想低通滤波，滤除0.5%的总功率，与原图接近说明边缘信息在0.5%以上的功率中

频率域图像增强



频率域函数H(u, v)

模糊且半径为5的ILPF

f(x, y)由黑色背景

下5个明亮的像素组成，

明亮点可看作冲激

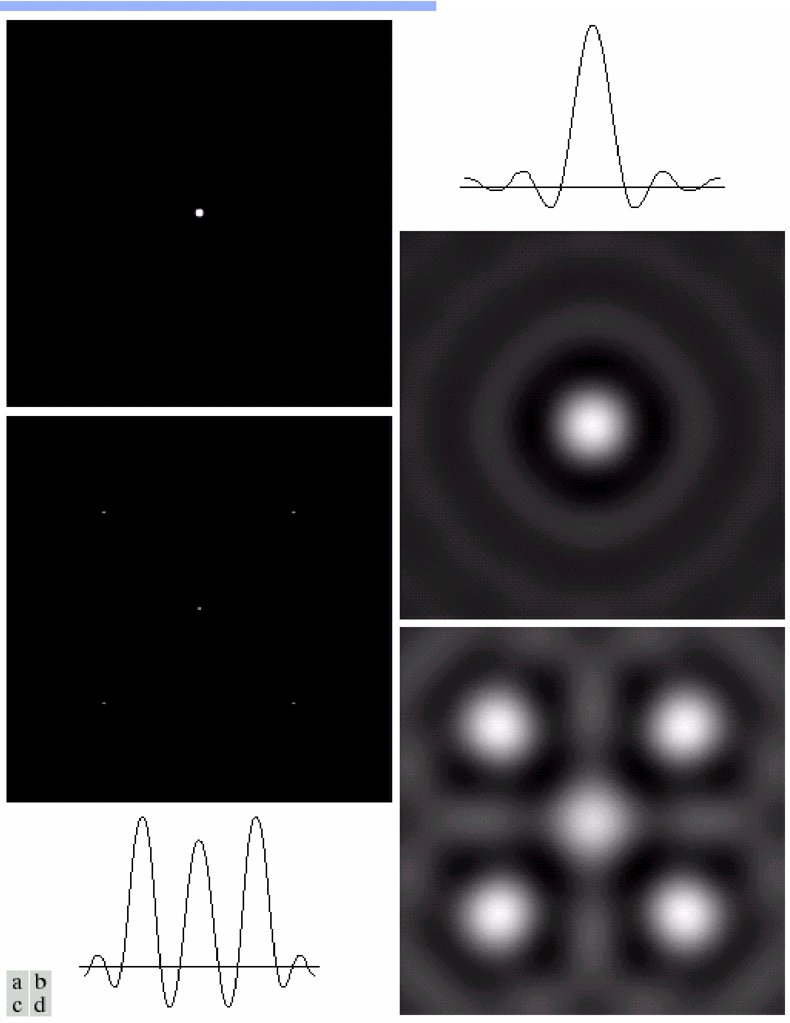


FIGURE 4.13 (a) A frequency-domain ILPF of radius 5. (b) Corresponding spatial filter (note the ringing). (c) Five impulses in the spatial domain, simulating the values of five pixels. (d) Convolution of (b) and (c) in the spatial domain.

对应空间域h(x, y)

中心开始的圆环周期

f(x, y)*h(x, y)，在每

个冲激处复制h(x, y)

的过程，振铃现象

频率域图像增强



- 巴特沃思低通滤波器

✓ n级巴特沃思低通滤波器(BLPF)定义如下

$$H(u,v)=\frac{1}{1+\left[D(u,v)/D_0\right]^{2n}}$$

D_0为截至频率距原点的距离，D(u, v)是点(u, v)距原点的距离

✓ 不同于ILPF，BLPF变换函数在通带与被滤除的频率之间没有明显的截断

✓ 当D(u, v)=D_0时，H(u, v)=0.5(最大值是1，当D(u, v)=0)

巴特沃思低通滤波器



透视图　　滤波器　　阶数从1到4的滤波器横截面

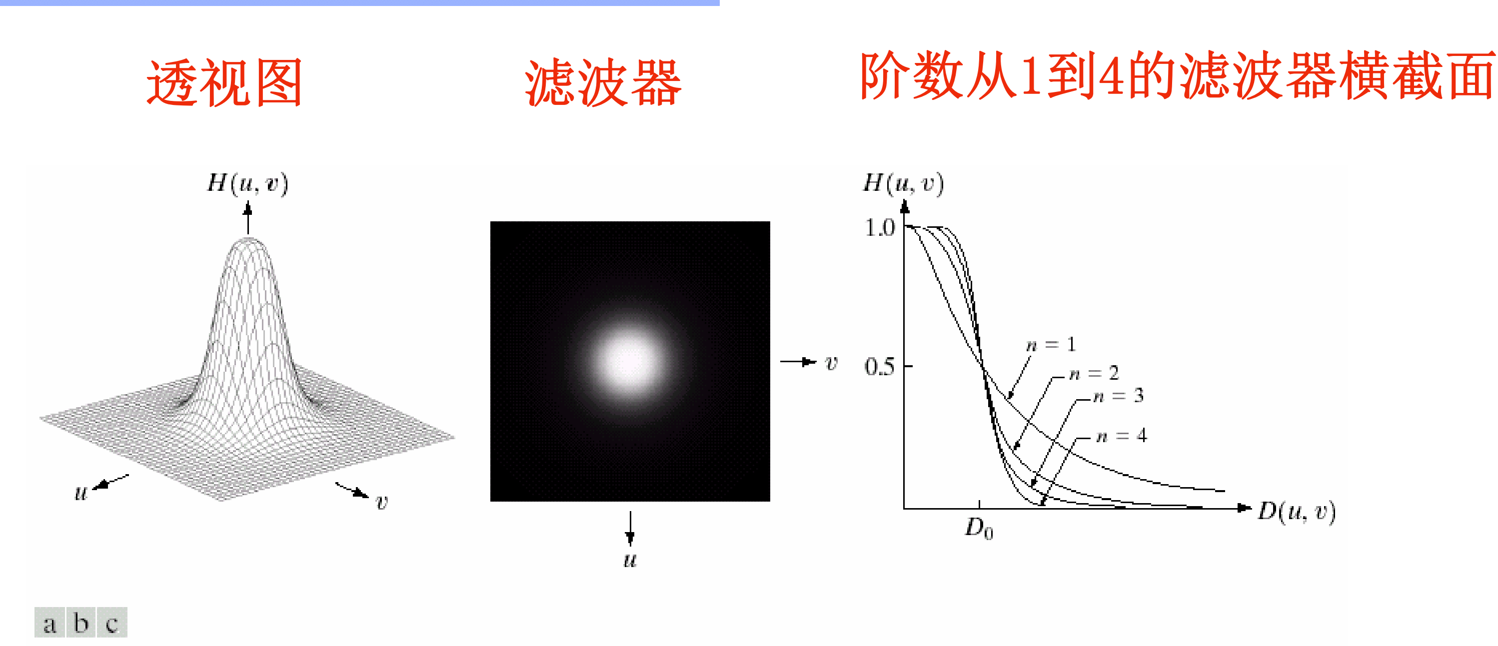


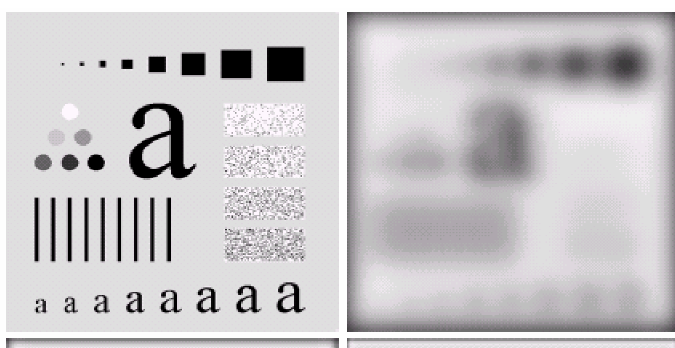
FIGURE 4.14 (a) Perspective plot of a Butterworth lowpass filter transfer function. (b) Filter displayed as an image. (c) Filter radial cross sections of orders 1 through 4.

应用：可用于平滑处理，如图像由于量化不足产生虚假轮廓时，常可用低通滤波进行平滑以改进图像质量。通常，BLPF的平滑效果好于ILPF（振铃现象）。

巴特沃思低通滤波器 $n = 2$

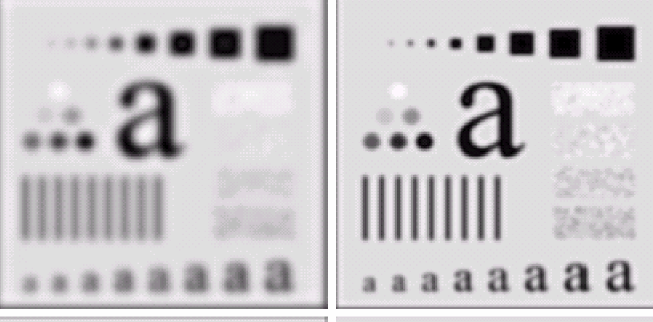


原图



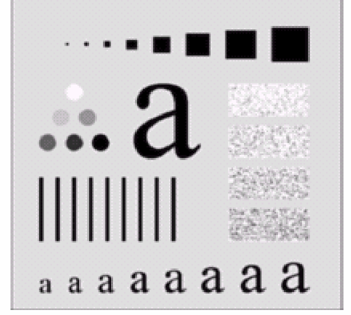
半径是5的BLPF滤波

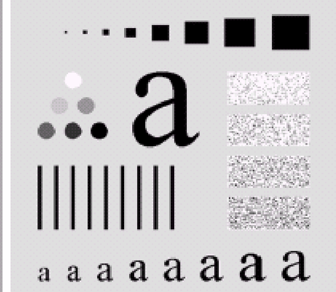
半径是15的BLPF滤波



半径是30的BLPF滤波

半径是80的BLPF滤波





半径是230的BLPF滤波

a b
c d
e f

FIGURE 4.15 (a) Original image. (b)–(f) Results of filtering with BLPFs of order 2, with cutoff frequencies at radii of 5, 15, 30, 80, and 230, as shown in Fig. 4.11(b). Compare with Fig. 4.12.

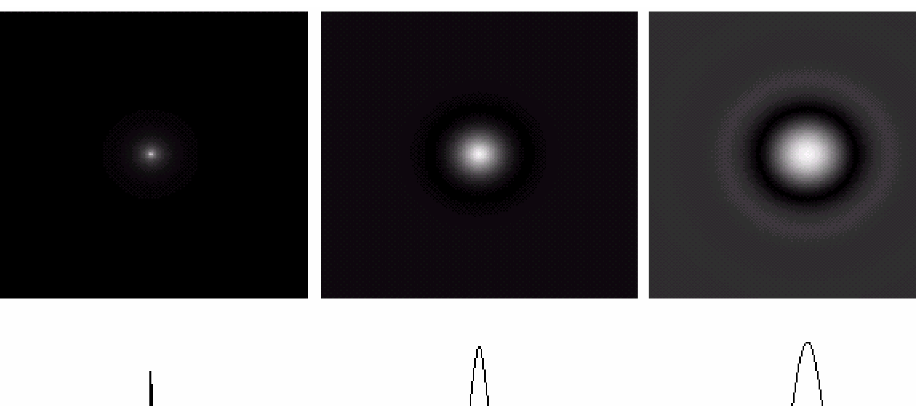
巴特沃思低通滤波器

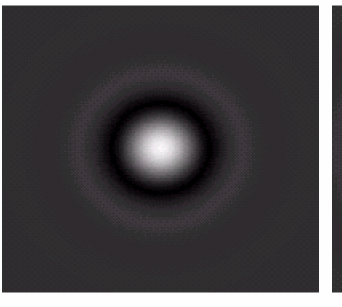


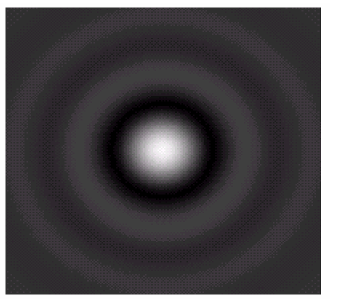
ILPF

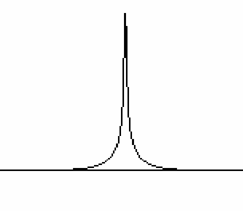
所有的滤波器都有半径为5的截至频率D_0

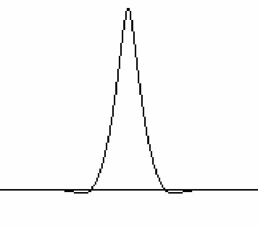
阶数n=1	阶数n=2	阶数n=5	阶数n=20
无振铃和负值	轻微振铃和负值	明显振铃和负值	与ILPF相似

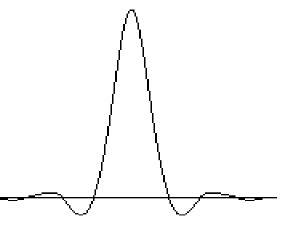


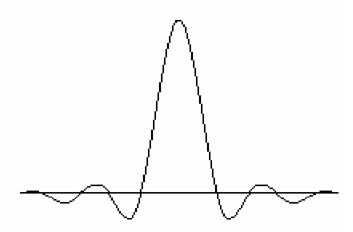












a b c d

FIGURE 4.16 (a)–(d) Spatial representation of BLPFs of order 1, 2, 5, and 20, and corresponding gray-level profiles through the center of the filters (all filters have a cutoff frequency of 5). Note that ringing increases as a function of filter order.

注：二阶BLPF处于有效低通滤波和可接受的振铃特征之间

频率域图像增强



- 高斯低通滤波器

  ✓ 二维高斯低通滤波器(GLPF)定义如下

$$H(u,v)=e^{-D(u,v)^2/2\sigma^2}$$

  ✓ D(u, v)是点(u, v)距原点的距离，使 $\sigma = D_0$

$$H(u,v)=e^{-D(u,v)^2/2D_0^2}$$

  ✓ 当D(u, v)=D_0时，滤波器下降到它最大值的0.607处

高斯低通滤波器



透视图　　滤波器　　各种D_0值的滤波器横截面

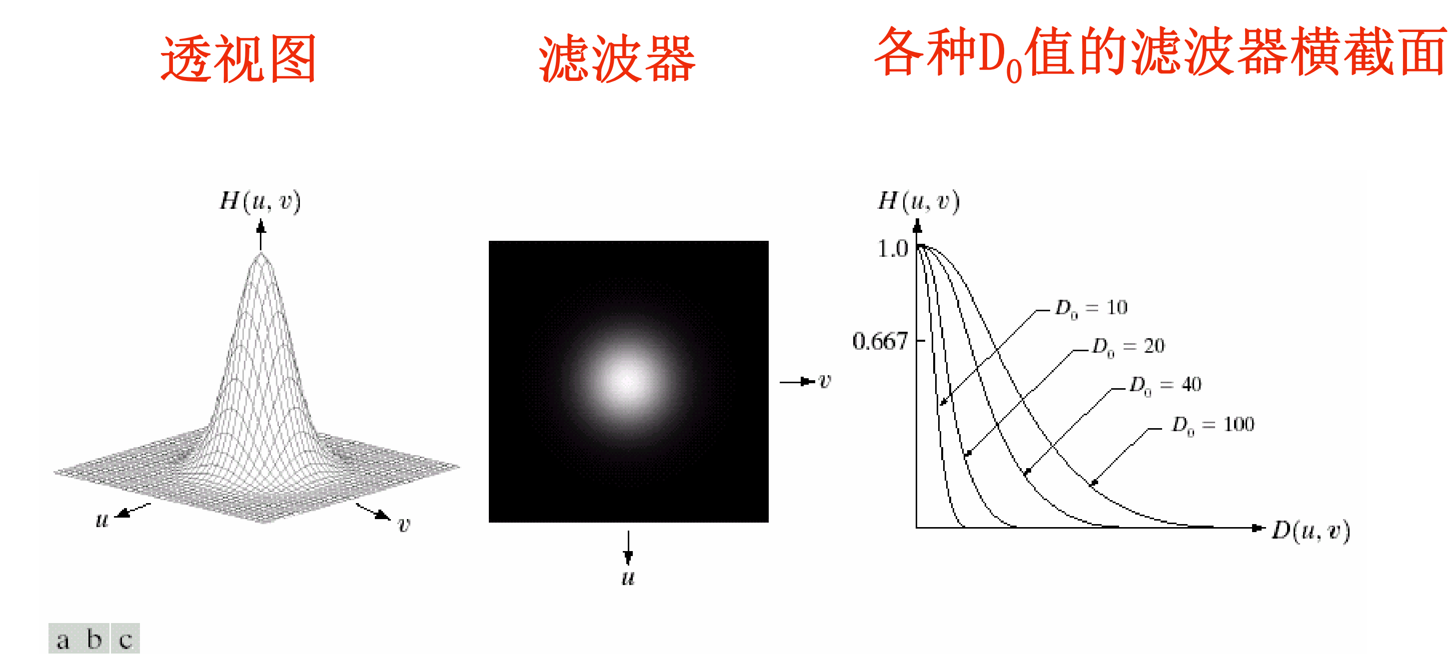


FIGURE 4.17 (a) Perspective plot of a GLPF transfer function. (b) Filter displayed as an image. (c) Filter radial cross sections for various values of D_0.

高斯低通滤波器

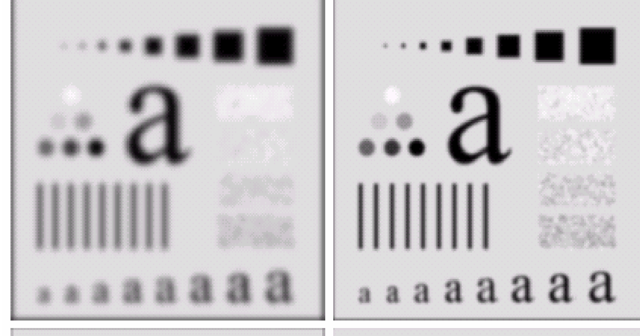


原图



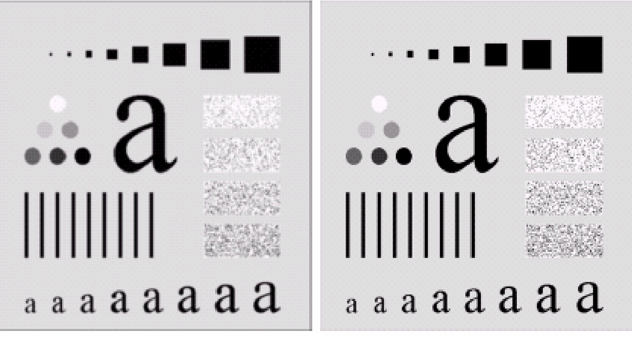
半径是5的GLPF滤波

半径是15的GLPF滤波



半径是30的GLPF滤波

半径是80的GLPF滤波



半径是230的GLPF滤波

FIGURE 4.18 (a) Original image. (b)–(f) Results of filtering with Gaussian lowpass filters with cutoff frequencies set at radii values of 5, 15, 30, 80, and 230, as shown in Fig. 4.11(b). Compare with Figs. 4.12 and 4.15.

a b
c d
e f

频率域图像增强



- 结论
  - ✓ GLPF不能达到有相同截止频率的二阶BLPF的平滑效果
  - ✓ GLPF没有振铃
  - ✓ 如果需要严格控制低频和高频之间截至频率的过渡，选用BLPF，代价是可能产生振铃

频率域图像增强



- 低通滤波器的应用实例：模糊, 平滑等
  - ✓ 字符识别：通过模糊图像，桥接断裂字符的裂缝
  - ✓ 印刷和出版业：从一幅尖锐的原始图像产生平滑、柔和的外观，如人脸，减少皮肤细纹的锐化程度和小斑点
  - ✓ 处理卫星和航空图像：尽可能模糊细节，而保留大的可识别特征。低通滤波通过消除不重要的特征来简化感兴趣特征的分析

字符识别举例



如打印、传真、复印文本等，字符失真、字符断裂等

D_0=80的高斯低通滤波器修复字符

a b

FIGURE 4.19 (a) Sample text of poor resolution (note broken characters in magnified view). (b) Result of filtering with a GLPF (broken character segments were joined).

Historically, certain computer programs were written using only two digits rather than four to define the applicable year. Accordingly, the company's software may recognize a date using "00" as 1900 rather than the year 2000.

Historically, certain computer programs were written using only two digits rather than four to define the applicable year. Accordingly, the company's software may recognize a date using "00" as 1900 rather than the year 2000.

用于机器识别系统识别断裂字符的预处理

人脸图像处理





D_0=100的GLPF滤波，细纹减少

D_0=80的GLPF滤波，细纹减少







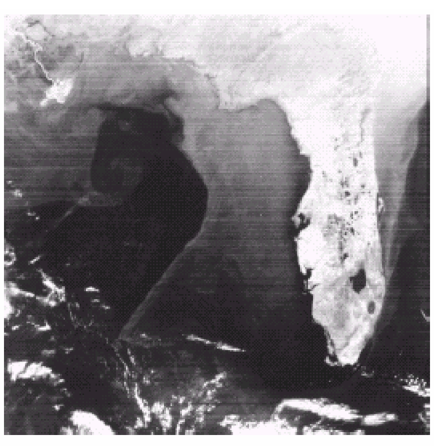
a b c

FIGURE 4.20 (a) Original image (1028 × 732 pixels). (b) Result of filtering with a GLPF with $D_0 = 100$. (c) Result of filtering with a GLPF with $D_0 = 80$. Note reduction in skin fine lines in the magnified sections of (b) and (c).

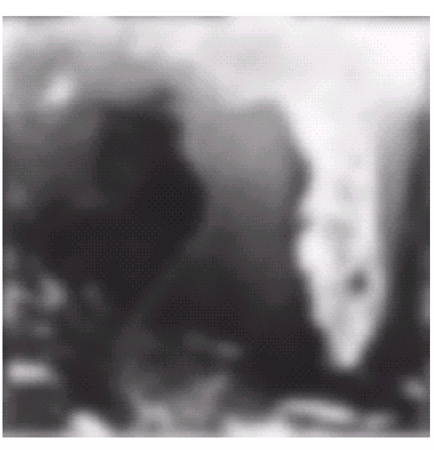
人脸图像处理



原图像　　D_0=30的GLPF滤波　　D_0=10的GLPF滤波，







a b c

FIGURE 4.21 (a) Image showing prominent scan lines. (b) Result of using a GLPF with $D_0 = 30$. (c) Result of using a GLPF with $D_0 = 10$. (Original image courtesy of NOAA.)

目的：尽可能模糊细节，而保留大的可识别特征

频率域图像增强



- 频率域锐化滤波器
  - ✓ 理想高通滤波器
  - ✓ 巴特沃思高通滤波器
  - ✓ 高斯高通滤波器
  - ✓ 频率域的拉普拉斯算子
  - ✓ 钝化模板、高频提升滤波和高频加强滤波

频率域锐化滤波器



透视图　　图像表示　　横截面

理想高通滤波器

巴特沃思高通滤波器

高斯高通滤波器

FIGURE 4.22 Top row: Perspective plot, image representation, and cross section of a typical ideal highpass filter. Middle and bottom rows: The same sequence for typical Butterworth and Gaussian highpass filters.

巴特沃思滤波器为理想滤波器的尖锐化和高斯滤波器的完全光滑之间的一种过渡

频率域锐化滤波器



理想高通滤波器　　巴特沃思高通滤波器　　高斯高通滤波器

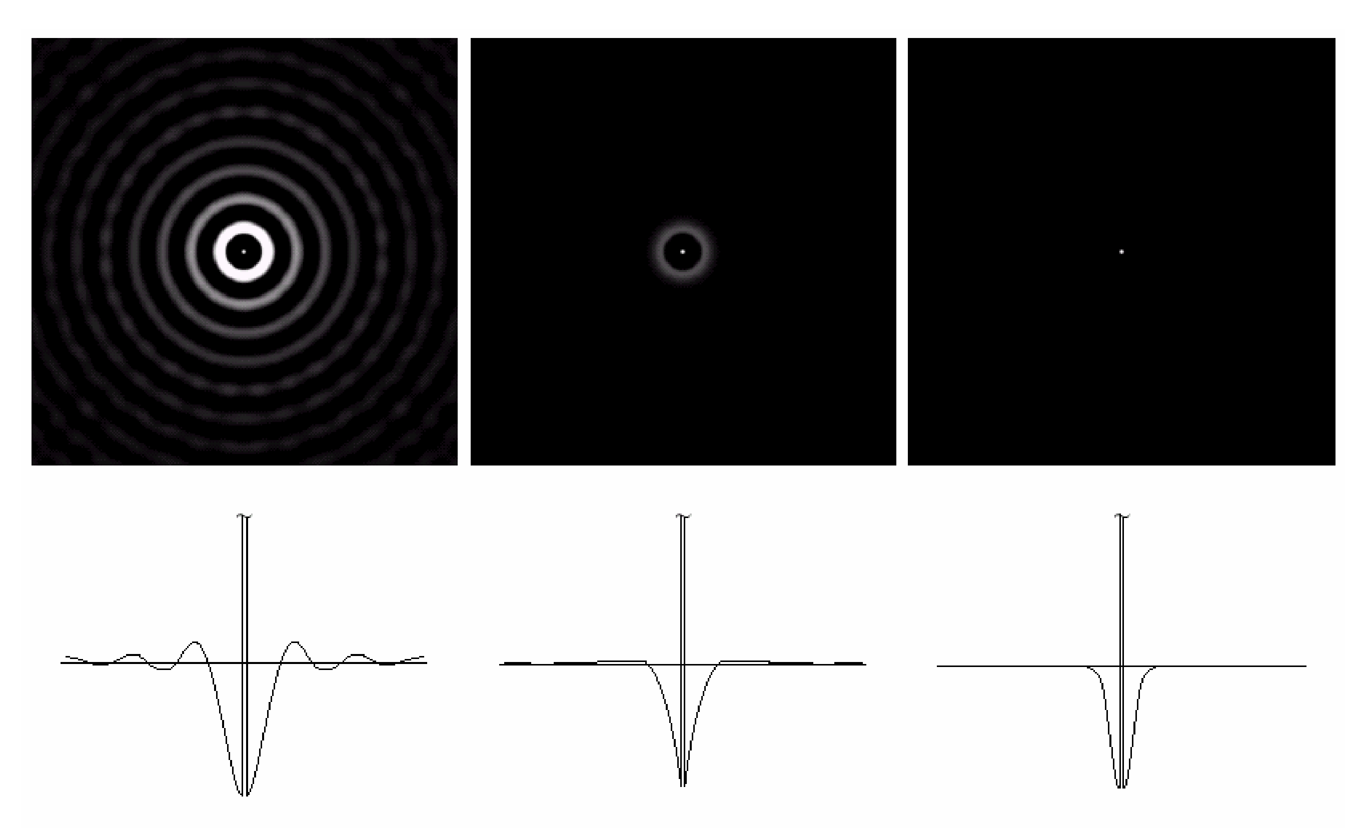


FIGURE 4.23 Spatial representations of typical (a) ideal, (b) Butterworth, and (c) Gaussian frequency domain highpass filters, and corresponding gray-level profiles.

频率域图像增强



- 理想高通滤波器

  ✓ 截断傅里叶变换中的所有低频成分，这些低频成分处于指定距离D_0之内

$$H(u,v)=\begin{cases}0 & D(u,v)\le D_0\\ 1 & D(u,v)>D_0\end{cases}$$

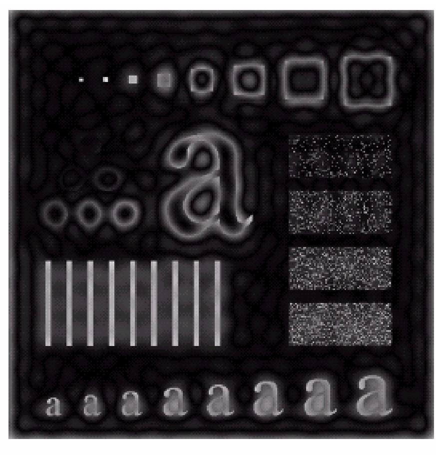
  ✓ 频率矩形的中心在(u, v)=(M/2, N/2)，从点(u, v)到中心（原点）的距离如下

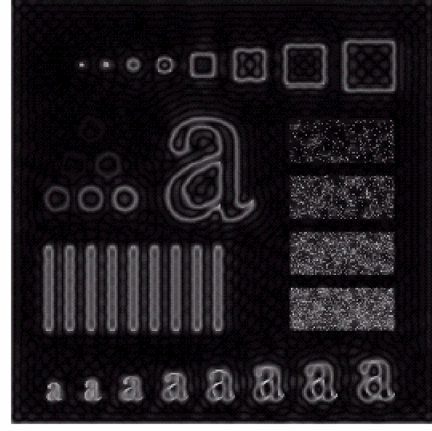
$$D(u,v)=\left[\left(u-\frac{M}{2}\right)^2+\left(v-\frac{N}{2}\right)^2\right]^{\frac{1}{2}}$$

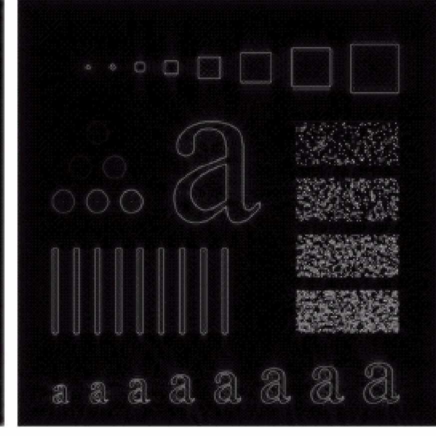
理想高通滤波器



图a：D_0=15　　图b：D_0=30　　图c：D_0=80







a b c

FIGURE 4.24 Results of ideal highpass filtering the image in Fig. 4.11(a) with D_0 = 15, 30, and 80, respectively. Problems with ringing are quite evident in (a) and (b).

结论：图a和b的振铃问题十分明显

频率域图像增强



- 巴特沃思高通滤波器
  - ✓ n阶且截至频率距原点的距离为D_0的巴特沃思高通滤波器(BHPF)定义为

$$H(u,v)=\frac{1}{1+\left[D_0/D(u,v)\right]^{2n}}$$

  - ✓ 推导

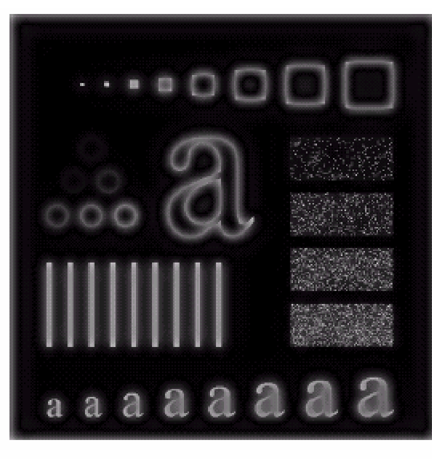
$$H(u,v)=1-H_{lp}(u,v)=1-\frac{1}{1+\left[D(u,v)/D_0\right]^{2n}}$$

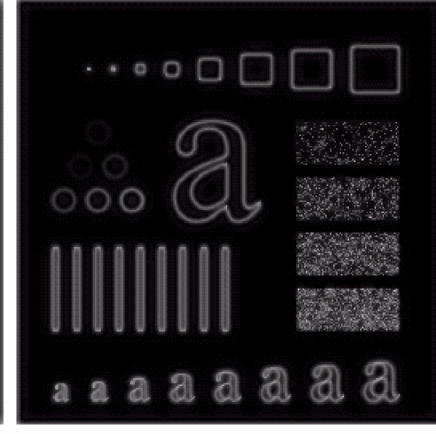
$$=\frac{-\left[D(u,v)/D_0\right]^{2n}}{1+\left[D(u,v)/D_0\right]^{2n}}=\frac{-1}{1+\left[D_0/D(u,v)\right]^{2n}}$$

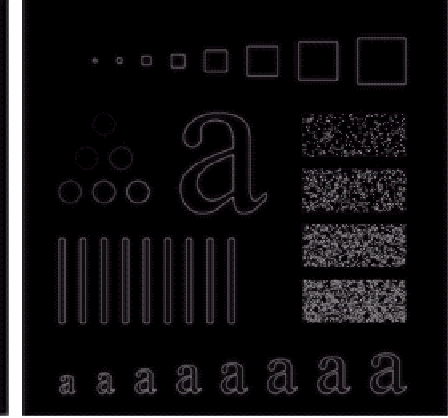
二阶巴特沃思高通滤波器



图a：D_0=15 图b：D_0=30 图c：D_0=80







a b c

FIGURE 4.25 Results of highpass filtering the image in Fig. 4.11(a) using a BHPF of order 2 with $D_0 = 15$, 30, and 80, respectively. These results are much smoother than those obtained with an ILPF.

结论：BHPF的结果比IHPF的结果平滑得多

频率域图像增强



- 高斯高通滤波器
  - ✓ 截频距原点为D_0的高斯高通滤波器(GHPF)定义为

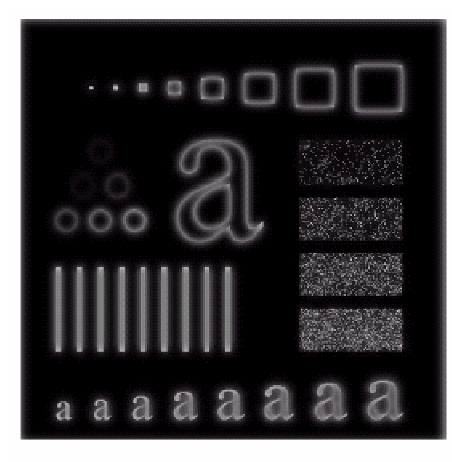
$$H(u,v) = 1 - e^{-D(u,v)^2 / 2D_0^2}$$

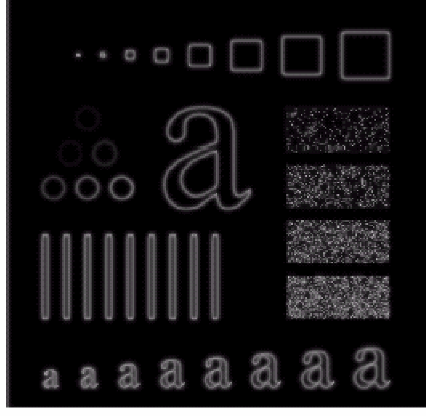
$$D(u,v) = \left[\left(u - \frac{M}{2}\right)^2 + \left(v - \frac{N}{2}\right)^2\right]^{\frac{1}{2}}$$

高斯高通滤波器



图a：D_0=15　　图b：D_0=30　　图c：D_0=80







a b c

FIGURE 4.26 Results of highpass filtering the image of Fig. 4.11(a) using a GHPF of order 2 with $D_0 = 15$, 30, and 80, respectively. Compare with Figs. 4.24 and 4.25.

结论：GHPF的结果比BHBF和IHPF的结果更平滑

频率域图像增强



- 频率域的拉普拉斯算子定义

$$\Im\left[\frac{d^n f(x)}{dx^n}\right]=(ju)^n F(u)$$

可以得到

$$\Im\left[\frac{\partial^2 f(x,y)}{\partial x^2}+\frac{\partial^2 f(x,y)}{\partial y^2}\right]=(ju)^2 F(u,v)+(jv)^2 F(u,v)$$

$$=-\left(u^2+v^2\right)F(u,v)$$

即

$$\Im\left[\nabla^2 f(x,y)\right]=-\left(u^2+v^2\right)F(u,v)=H(u,v)F(u,v)$$

所以 $H(u,v)=-\left(u^2+v^2\right)$

频率域图像增强



- 原点从(0,0)移到(M/2,N/2)，所以，滤波函数平移为

$$H(u,v)=-\left[(u-M/2)^2+(v-N/2)^2\right]$$

空间域拉普拉斯算子过滤后的图像可由计算H(u,v)F(u,v)的反傅里叶变换得到

$$\nabla^2 f(x,y)=\Im^{-1}\left\{-\left[(u-M/2)^2+(v-N/2)^2\right]F(u,v)\right\}$$

傅里叶变换对表示空间域拉普拉斯算子和频率域的双向关系

$$\nabla^2 f(x,y)\Leftrightarrow -\left[(u-M/2)^2+(v-N/2)^2\right]F(u,v)$$

频率域图像增强



图a:频率域拉普拉斯的三维图

图b:图a的图像表示

图c:从图b的傅里叶反变换得到的空间域拉普拉斯

图d：图c的放大图像

图e：通过图d中心的灰度剖面图

0	1	0
1	–4	1
0	1	0

图f：空间域的拉普拉斯模板

a b
c d e
f

FIGURE 4.27 (a) 3-D plot of Laplacian in the frequency domain. (b) Image representation of (a). (c) Laplacian in the spatial domain obtained from the inverse DFT of (b). (d) Zoomed section of the origin of (c). (e) Gray-level profile through the center of (d). (f) Laplacian mask used in Section 3.7.

频率域图像增强



- 从原始图像中减去拉普拉斯算子部分，形成g(x, y)的增强图像

$$g(x,y) = f(x,y) - \nabla^2 f(x,y)$$

拉普拉斯举例说明



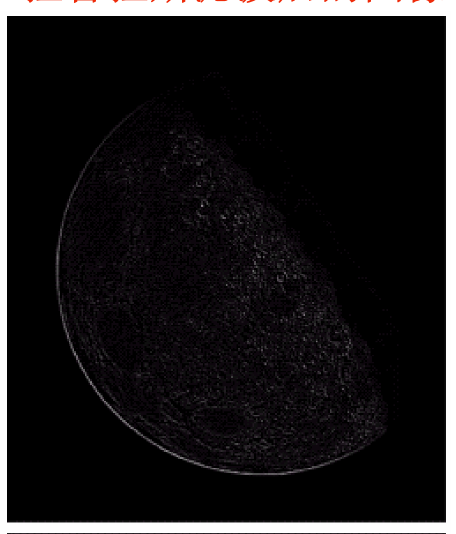
a b
c d

FIGURE 4.28
(a) Image of the North Pole of the moon.
(b) Laplacian filtered image.
(c) Laplacian image scaled.
(d) Image enhanced by using Eq. (4.4-12).
(Original image courtesy of NASA.)

月球北极图像



拉普拉斯滤波后的图像



标定后的图像：

因为具有正值

和负值





增强的图像

原图＋拉普拉斯图像

频率域图像增强



- 频率域锐化滤波器
  - ✓ 理想高通滤波器
  - ✓ 巴特沃思高通滤波器
  - ✓ 高斯高通滤波器
  - ✓ 频率域的拉普拉斯算子
  - ✓ 钝化模板、高频提升滤波和高频加强滤波

频率域图像增强



- 为什么要进行高频提升和高频加强？
  - ✓ 高频滤波后的图像，其背景平均强度减小到接近黑色（因为高通滤波器滤除了傅里叶变换的零频率成分：$F(0,0)=\overline{f}(x,y)=0$）
  - ✓ 解决办法：把原始图像加到过滤后的结果，如拉普拉斯算子增强，这种处理称为高频提升过滤

频率域图像增强



- 钝化模板(锐化或高通图像)：从一幅图像减去其自身模糊图像而生成的锐化图像构成。在频率域，即从图像本身减去低通滤波（模糊）后的图像而得到高通滤波（锐化）的图像

$$f_{hp}(x,y) = f(x,y) - f_{lp}(x,y)$$

- 高频提升过滤：

$$\begin{aligned} f_{hp}(x,y) &= Af(x,y) - f_{lp}(x,y) \\ &= (A-1)f(x,y) + f(x,y) - f_{lp}(x,y) \\ &= (A-1)f(x,y) + f_{hp}(x,y) \end{aligned}$$

当A=1,即高通过滤；当A>1,累加图像本身

频率域图像增强



- 高频提升过滤（续）：

由 $$H_{hp}(u,v)=1-H_{lp}(u,v)$$

高频提升过滤可以定义为

$$H_{hp}(u,v)=(A-1)+H_{hp}(u,v)$$

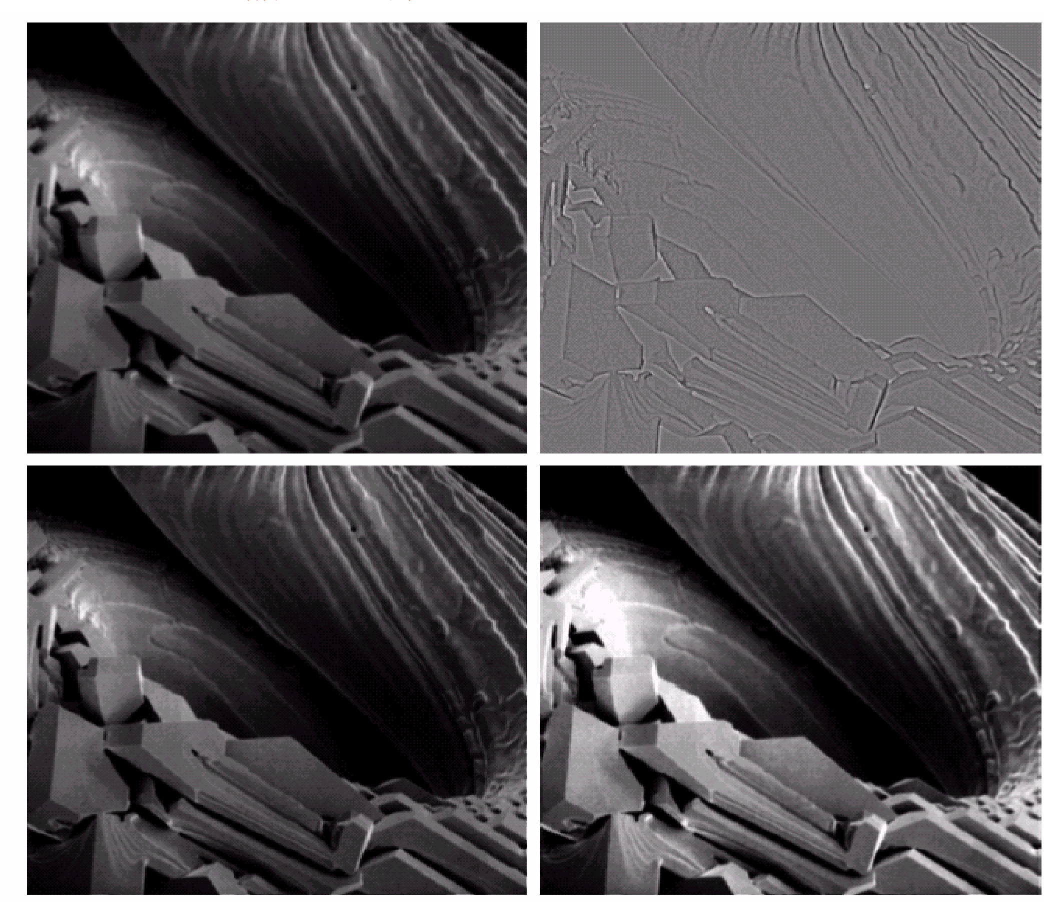
高频提升过滤举例——频率域滤波



图a：输入图像　　　图b：图a的拉普拉斯图像

a	b
c	d

FIGURE 4.29 Same as Fig. 3.43, but using frequency domain filtering. (a) Input image. (b) Laplacian of (a). (c) Image obtained using Eq. (4.4-17) with $A = 2$. (d) Same as (c), but with $A = 2.7$. (Original image courtesy of Mr. Michael Shaffer, Department of Geological Sciences, University of Oregon, Eugene.)



图c：A=2　　　图d：A=2.7

频率域图像增强



- 高频提升加强：

$$H_{hp}(u,v) = a + bH_{hp}(u,v)$$

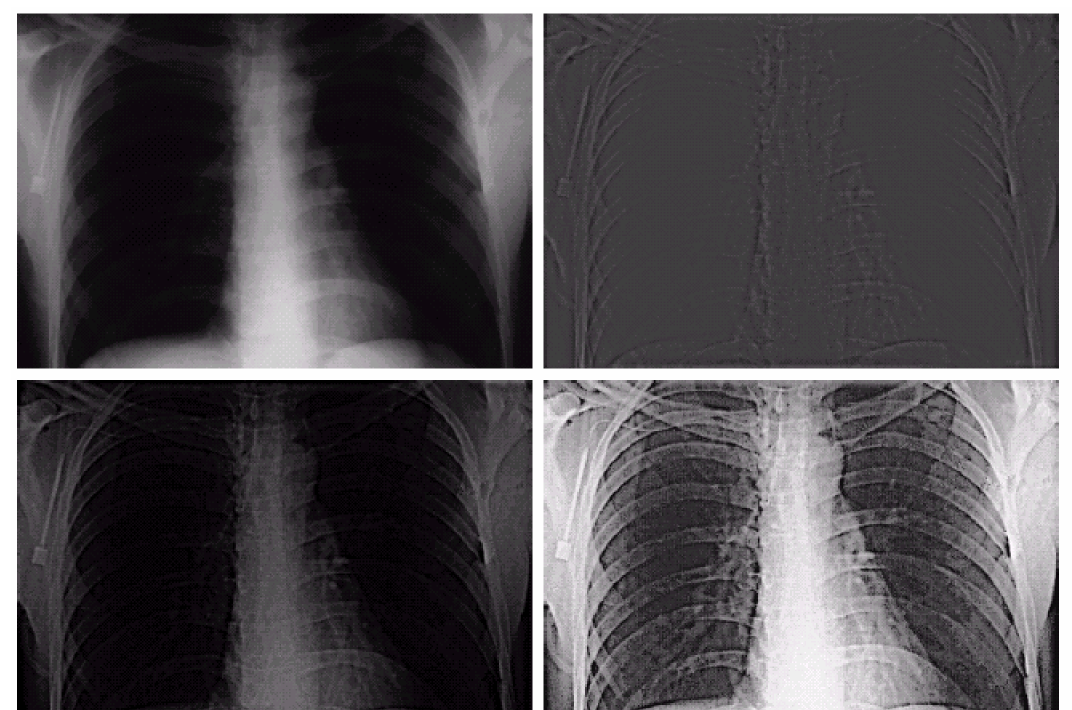
$$a \geq 0, b > a$$

  - ✓ 用图像的高频成分进行增强
  - ✓ 增加a的目的是使零频率不被滤波器过滤
  - ✓ 当a=A-1, b=1时转化为高频提升过滤
  - ✓ 当b>1，高频得到加强

高频提升加强举例



图a：胸部X光图像　　　图b：巴特沃思高通滤波的结果



a b
c d

FIGURE 4.30
(a) A chest X-ray image. (b) Result of Butterworth highpass filtering. (c) Result of high-frequency emphasis filtering. (d) Result of performing histogram equalization on (c). (Original image courtesy Dr. Thomas R. Gest, Division of Anatomical Sciences, University of Michigan Medical School.)

图c：高频增强滤波的结果　　　图d：图c直方图均衡化的结果

注：高频加强 ＋ 直方图均衡化的混合图像增强方法

数字图像处理（6）



任何问题？

研究生课程



数字图像处理

Digital Image Processing

彭宇新

北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

图像复原



- 图像退化/复原过程的模型
- 噪声模型
- 空间域滤波复原（唯一退化是噪声）
- 频率域滤波复原（削减周期噪声）

图像复原



- 什么是退化？

成像过程中的"退化"，是指由于成像系统各种因素的影响，使得图像质量降低

- 引起图像退化的原因
  - ✓ 成像系统的散焦
  - ✓ 成像设备与物体的相对运动
  - ✓ 成像器材的固有缺陷
  - ✓ 外部干扰等

图像复原



- 图像复原概述
  - ✓ 与图像增强相似，图像复原的目的也是改善图像质量
  - ✓ 图像增强主要是一个主观过程，而图像复原主要是一个客观过程
  - ✓ 图像增强被认为是一种对比度拉伸，提供给用户喜欢接收的图像；而图像复原技术追求恢复原始图像的最优估值
  - ✓ 图像复原技术可以使用空间域或频率域滤波器实现

图像复原



- 图像复原概述（续）

  ✓ 图像复原可以看作图像退化的逆过程，是将图像退化的过程加以估计，建立退化的数学模型后，补偿退化过程造成的失真

  ✓ 在图像退化确知的情况下，图像退化的逆过程是有可能进行的

  ✓ 但实际情况经常是退化过程并不知晓，这种复原称为盲目复原

  ✓ 由于图像模糊的同时，噪声和干扰也会同时存在，这也为复原带来了困难和不确定性

图像复原



- 图像退化/复原过程的模型

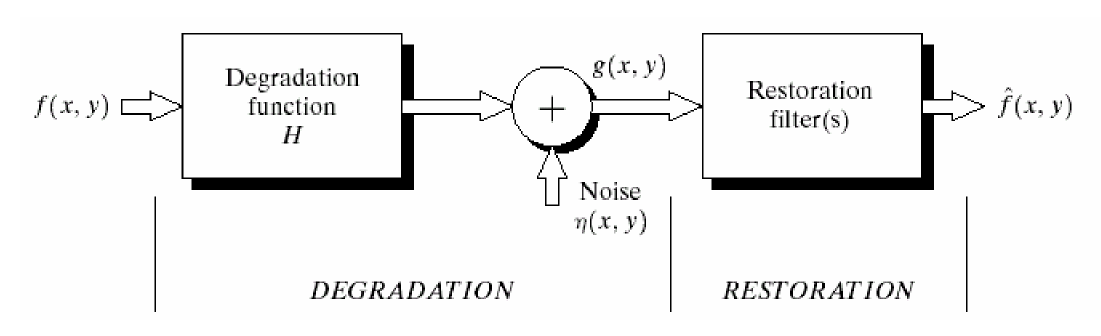


FIGURE 5.1 A model of the image degradation/restoration process.

- ✓ f(x, y)表示一幅输入图像
- ✓ g(x, y)是f(x, y)产生的一幅退化图像
- ✓ H表示退化函数
- ✓ $\eta(x, y)$表示外加噪声
- ✓ 给定g(x, y)，H和$\eta(x, y)$，怎样获得关于原始图像的近似估计$\hat{f}(x, y)$？

图像复原



- 如果系统H是一个线性、位置不变性的过程，退化图像可以表示为

$$g(x,y)=h(x,y)*f(x,y)+\eta(x,y)$$

⇩ 空间域上的卷积等同于频率域上的乘积

$$G(u,v)=H(u,v)F(u,v)+N(u,v)$$

h(x, y)表示退化函数的空间描述

图像复原



- 噪声模型
  - ✓ 数字图像的噪声主要来源于图像的获取和传输过程
    - 图像获取的数字化过程，如图像传感器的质量和环境条件
    - 图像传输过程中传输信道的噪声干扰，如通过无线网络传输的图像会受到光或其它大气因素的干扰

图像复原

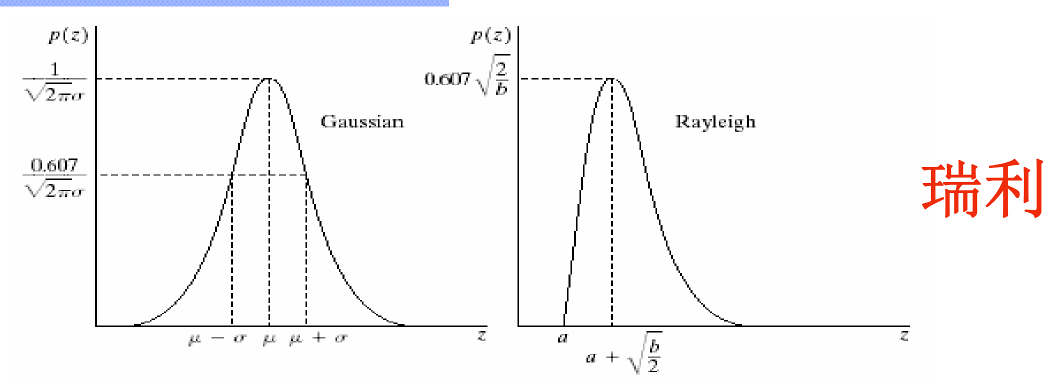


- 一些重要的噪声
  - ✓ 高斯噪声
  - ✓ 瑞利噪声
  - ✓ 伽马（爱尔兰）噪声
  - ✓ 指数分布噪声
  - ✓ 均匀分布噪声
  - ✓ 脉冲噪声（椒盐噪声）

一些重要噪声的概率密度函数(PDF)

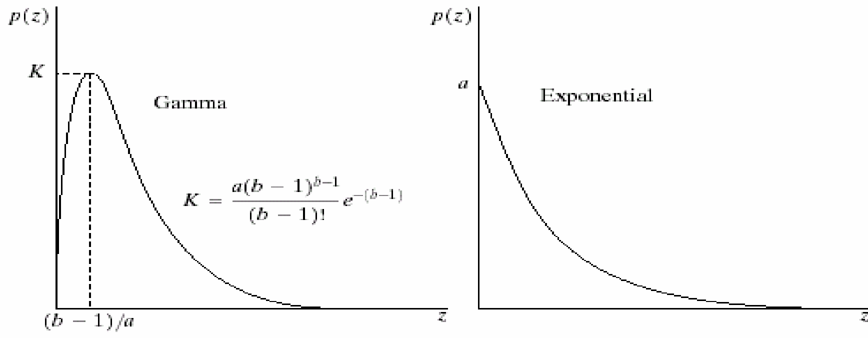






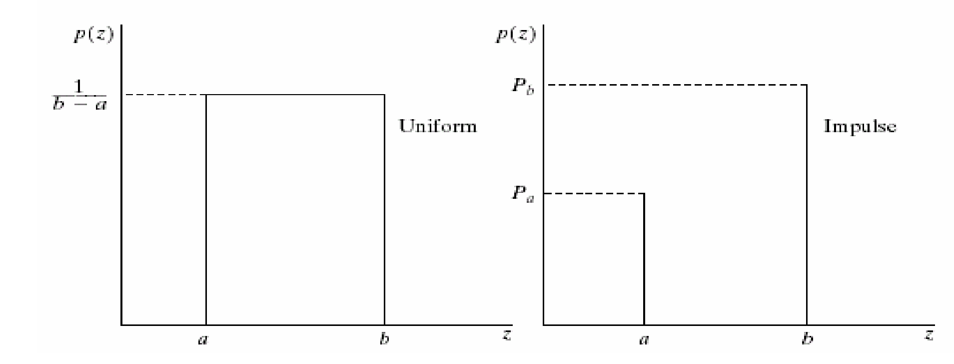
瑞利

伽马



指数





脉冲

图像复原



- 高斯噪声
  - ✓ 高斯噪声的概率密度函数(PDF)由下式给出

$$p(z)=\frac{1}{\sqrt{2\pi\sigma}}e^{-(z-\mu)^2/2\sigma^2}$$

  - ✓ z的值有70%落在 $[(\mu-\sigma),(\mu+\sigma)]$ 范围内，有95%落在 $[(\mu-2\sigma),(\mu+2\sigma)]$ 范围内

图像复原



- 瑞利噪声
  - ✓ 瑞利噪声的PDF由下式给出

$$p(z)=\begin{cases}\frac{2}{b}(z-a)e^{-(z-a)^2/b} & z \ge a \\ 0 & z < a\end{cases}$$

$$\mu = a + \sqrt{\pi b / 4}$$

$$\sigma^2 = \frac{b(4-\pi)}{4}$$

  - ✓ 距离原点的位移是a
  - ✓ 函数曲线向右变形

图像复原



- 伽马（爱尔兰）噪声
  - ✓ 伽马噪声的PDF由下式给出

$$p(z)=\begin{cases}\dfrac{a^b z^{b-1}}{(b-1)!}e^{-az} & z \ge a \\ 0 & z < a\end{cases}$$

$$\mu = \frac{b}{a}$$

$$\sigma^2 = \frac{b}{a^2}$$

图像复原



- 指数分布噪声
  - ✓ 指数噪声的PDF由下式给出，其中, a>0

$$p(z)=\begin{cases} ae^{-az} & z \geq a \\ 0 & z < a \end{cases}$$

$$\mu = \frac{1}{a}$$

$$\sigma^2 = \frac{1}{a^2}$$

  - ✓ 指数分布的PDF是当b=1时爱尔兰分布的特殊情况

图像复原



- 均匀分布噪声
  - ✓ 均匀分布噪声的PDF由下式给出

$$p(z)=\begin{cases}\dfrac{1}{b-a} & a\le z\le b\\ 0 & \text{其它}\end{cases}$$

$$\mu=\frac{a+b}{2}$$

$$\sigma^2=\frac{(b-a)^2}{12}$$

图像复原



- 脉冲噪声（椒盐噪声）
  - ✓ 脉冲噪声的PDF由下式给出

$$p(z)=\begin{cases} P_a & z=a \\ P_b & z=b \\ 0 & \text{其它} \end{cases}$$

  - ✓ 如果p_a或p_b为零，则脉冲噪声称为单极脉冲
  - ✓ 如果p_a或p_b均不为零，则脉冲噪声称为双极脉冲噪声或椒盐噪声

图像复原



- 脉冲噪声（椒盐噪声）（续）
  - ✓ 脉冲噪声可以为正，也可为负
  - ✓ 标定以后，脉冲噪声总是数字化为最大值（纯黑或纯白）
  - ✓ 通常，负脉冲以黑点（胡椒点）出现，正脉冲以白点（盐点）出现

图像复原



- 几种噪声的运用
  - ✓ 高斯噪声源于电子电路噪声和由低照明度或高温带来的传感器噪声
  - ✓ 瑞利噪声对分布在图像范围内特征化噪声有用
  - ✓ 伽马分布和指数分布用于激光成像噪声
  - ✓ 均匀密度分布作为模拟随机数产生器的基础
  - ✓ 脉冲噪声用于成像中的短暂停留中，如错误的开关操作

样本噪声图像和它们的直方图



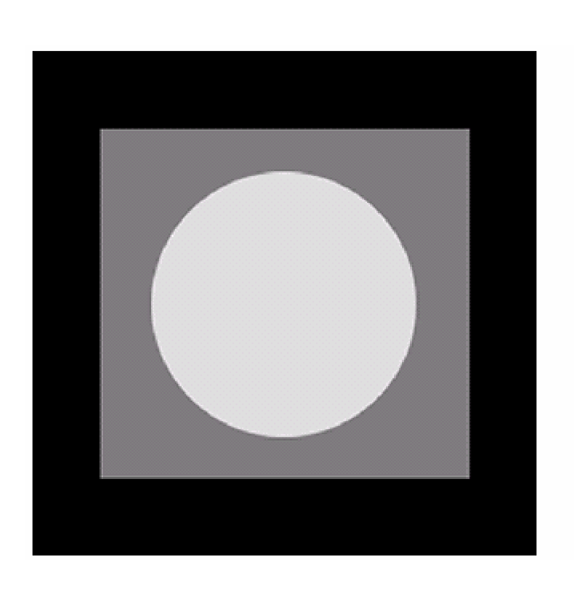


FIGURE 5.3 Test pattern used to illustrate the characteristics of the noise PDFs shown in Fig. 5.2.

- 用于噪声模型的测试图
- 由简单、恒定的区域组成
- 仅仅有3个灰度级的变化

样本噪声图像和它们的直方图



高斯噪声　瑞利噪声　伽马噪声

图像

直方图

Gaussian　Rayleigh　Gamma

a b c
d e f

FIGURE 5.4 Images and histograms resulting from adding Gaussian, Rayleigh, and gamma noise to the image in Fig. 5.3.

样本噪声图像和它们的直方图



指数噪声 均匀噪声 椒盐噪声

图像

直方图

Exponential Uniform Salt & Pepper

g h i
j k l

FIGURE 5.4 *(Continued)* Images and histograms resulting from adding exponential, uniform, and impulse noise to the image in Fig. 5.3.

样本噪声图像和它们的直方图



- 结论
  - 上述噪声图像的直方图和它们的概率密度函数曲线对应相似
  - 前面5种噪声的图像并没有显著不同
  - 但它们的直方图具有明显的区别

图像复原



- 周期噪声
  - ✓ 周期噪声是在图像获取中从电力或机电干扰中产生
  - ✓ 周期噪声可以通过频率域滤波显著减少

图像复原



- 空间域滤波复原（唯一退化是噪声）

当唯一退化是噪声时，

$$g(x,y) = f(x,y) + \eta(x,y)$$
$$G(u,v) = F(u,v) + N(u,v)$$

✓ 噪声项未知，不能从g(x, y)或G(u, v)减去噪声

✓ 可以选择空间滤波方法进行图像复原

图像复原



- 图像复原的空间滤波器
  - ✓ 均值滤波器

    算术均值滤波器、几何均值滤波器、谐波均值滤波器、逆谐波均值滤波器

  - ✓ 顺序统计滤波器

    中值滤波器、最大值滤波器、最小值滤波器、中点滤波器、修正后的阿尔法均值滤波器

  - ✓ 自适应滤波器

    自适应局部噪声消除滤波器、自适应中值滤波器

均值滤波器



- 算术均值滤波器

$$\hat{f}(x,y)=\frac{1}{mn}\sum_{(s,t)\in S_{xy}} g(s,t)$$

✓ S_{xy}表示中心在(x, y)，尺寸为m×n的矩形窗口

✓ 平滑了一幅图像的局部变化

✓ 在模糊了结果的同时减少了噪声

均值滤波器



- 几何均值滤波器

$$\hat{f}(x,y)=\left[\prod_{(s,t)\in S_{xy}} g(s,t)\right]^{\frac{1}{mn}}$$

✓ 几何均值滤波器所达到的平滑度可以与算术均值滤波器相比

✓ 但几何均值滤波器在滤波过程中，与算术均值滤波器相比，会丢失更少的图像细节——相对锐化

均值滤波器



- 谐波均值滤波器

$$\hat{f}(x,y)=\frac{mn}{\sum_{(s,t)\in S_{xy}}\frac{1}{g(s,t)}}$$

✓ 谐波均值滤波器对于“盐”噪声效果好，但不适用于“胡椒”噪声

✓ 善于处理高斯噪声等

均值滤波器



- 逆谐波均值滤波器

$$\hat{f}(x,y)=\frac{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q+1}}{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q}}$$

✓ Q称为滤波器的阶数。当Q为正数时，用于消除“胡椒”噪声；当Q为负数时，用于消除“盐”噪声，但不能同时消除“椒盐”噪声

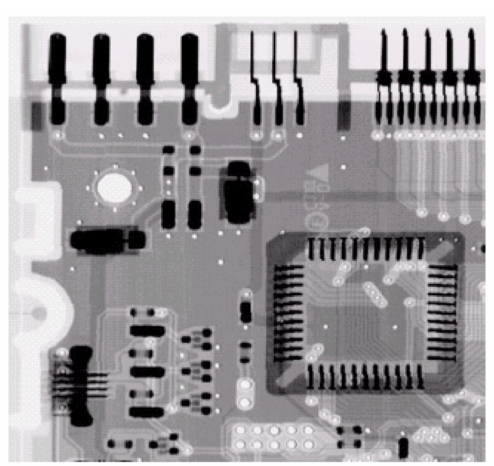
✓ 当Q=0，逆谐波均值滤波器转变为算术均值滤波器

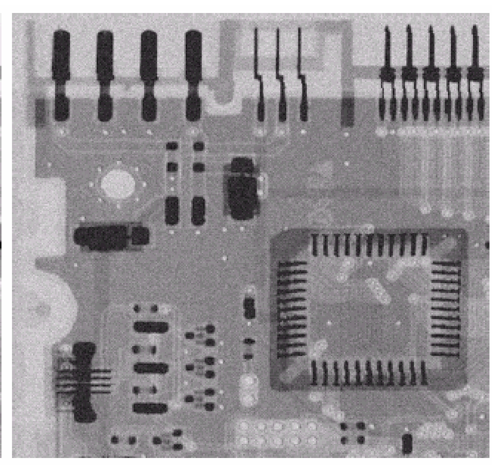
✓ 当Q=-1，逆谐波均值滤波器转变为谐波均值滤波器

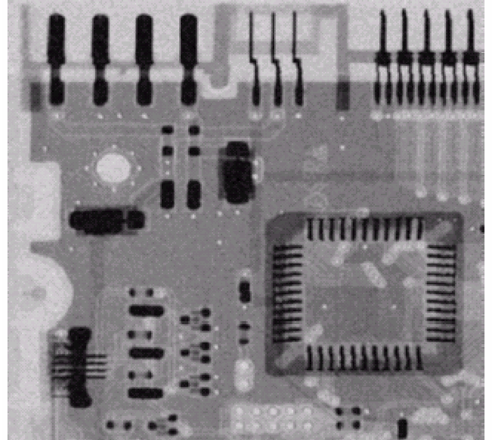
均值滤波举例

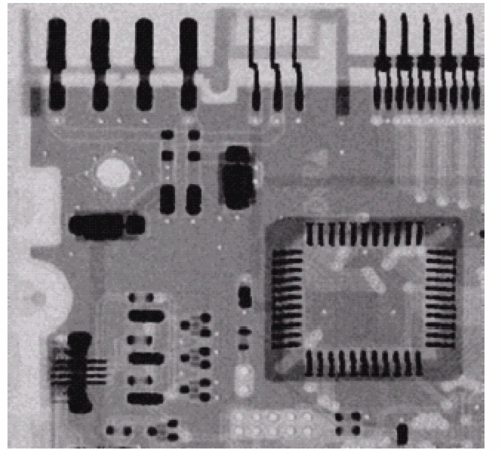


原图　　　被均值为0，方差为400的高斯噪声污染









3×3算术均值滤波器　　3×3几何均值滤波器，图像更清晰

a b
c d

FIGURE 5.7 (a) X-ray image. (b) Image corrupted by additive Gaussian noise. (c) Result of filtering with an arithmetic mean filter of size 3 × 3. (d) Result of filtering with a geometric mean filter of the same size. (Original image courtesy of Mr. Joseph E. Pascente, Lixi, Inc.)

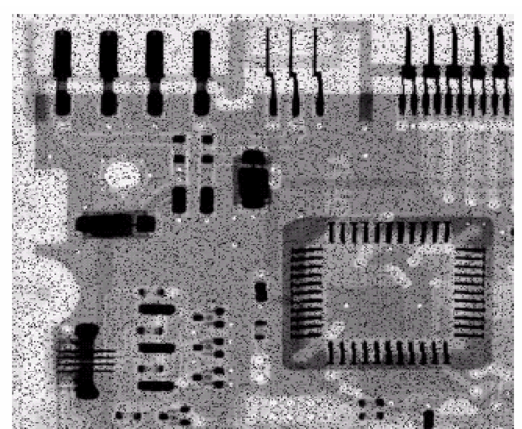
均值滤波举例

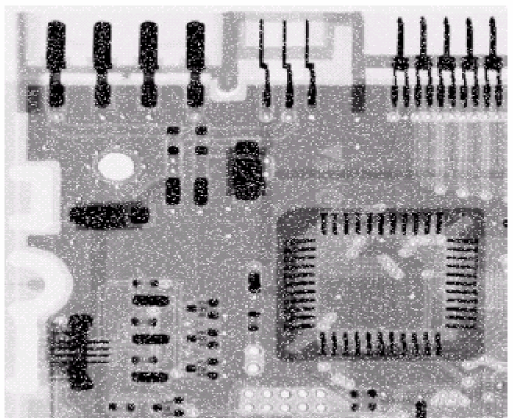


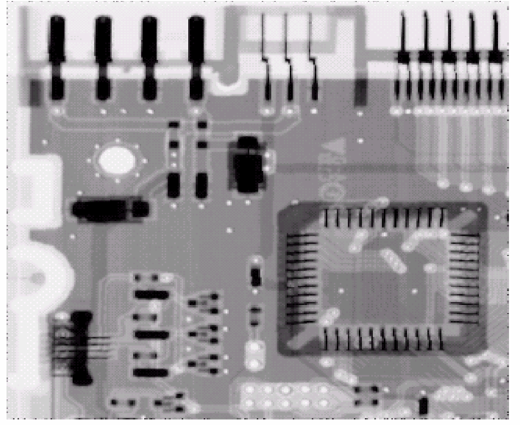
“胡椒”噪声干扰图像 “盐”噪声干扰图像

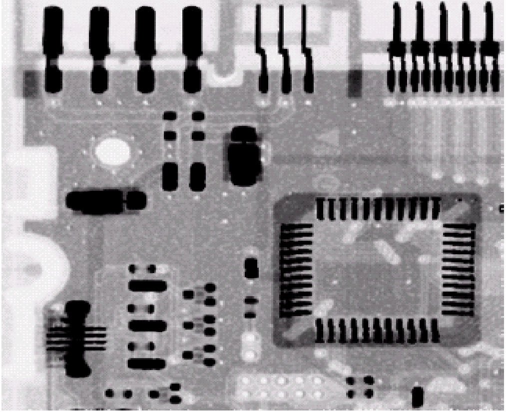
a b
c d

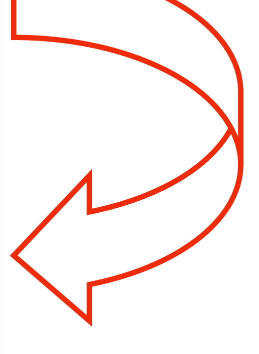
FIGURE 5.8 (a) Image corrupted by pepper noise with a probability of 0.1. (b) Image corrupted by salt noise with the same probability. (c) Result of filtering (a) with a 3 × 3 contraharmonic filter of order 1.5. (d) Result of filtering (b) with $Q = -1.5$.

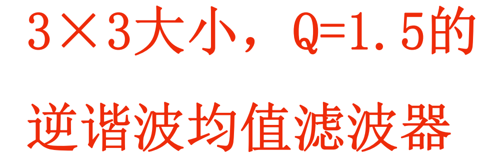














均值滤波器



- 总结
  - 算术均值滤波器和几何均值滤波器适合于处理高斯或均匀等随机噪声
  - 谐波均值滤波器适合于处理脉冲噪声

    缺点：必须事先知道噪声是暗噪声还是亮噪声，以便于选择合适的Q符号

顺序统计滤波器



- 中值滤波器

$$\hat{f}(x,y) = \underset{(s,t)\in S_{xy}}{median}\{g(s,t)\}$$

✓ 在相同尺寸下，比起均值滤波器引起的模糊少

✓ 对单极或双极脉冲噪声非常有效

顺序统计滤波器



- 最大值滤波器

$$\hat{f}(x,y) = \max_{(s,t)\in S_{xy}} \{g(s,t)\}$$

✓ 用于发现图像中的最亮点

✓ 可以有效过滤“胡椒”噪声（因为“胡椒”噪声是非常低的值）

顺序统计滤波器



- 最小值滤波器

$$\hat{f}(x,y) = \min_{(s,t)\in S_{xy}} \{g(s,t)\}$$

✓ 用于发现图像中的最暗点

✓ 可以有效过滤“盐”噪声（因为“盐”噪声是非常高的值）

顺序统计滤波器



- 中点滤波器

$$\hat{f}(x,y)=\frac{1}{2}\left[\max_{(s,t)\in S_{xy}}\{g(s,t)\}+\min_{(s,t)\in S_{xy}}\{g(s,t)\}\right]$$

✓ 结合了顺序统计和求平均

✓ 对于高斯和均匀随机分布这类噪声有最好的效果

顺序统计滤波器



- 修正后的阿尔法均值滤波器

$$\hat{f}(x,y)=\frac{1}{mn-d}\sum_{(s,t)\in S_{xy}} g_r(s,t)$$

✓ 在S_{xy}邻域内去掉g(s, t)最高灰度值的d/2和最低灰度值的d/2

✓ $g_r(s,t)$代表剩余的mn-d个像素

✓ 当d=0，退变为算术均值滤波器

✓ 当d=(mn-1)/2，退变为中值滤波器

✓ 当d取其它值时，适用于包括多种噪声的情况下，例如高斯噪声和椒盐噪声混合的情况

中值滤波器举例

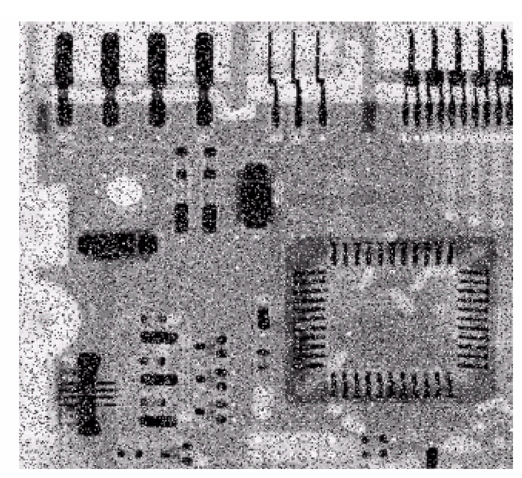


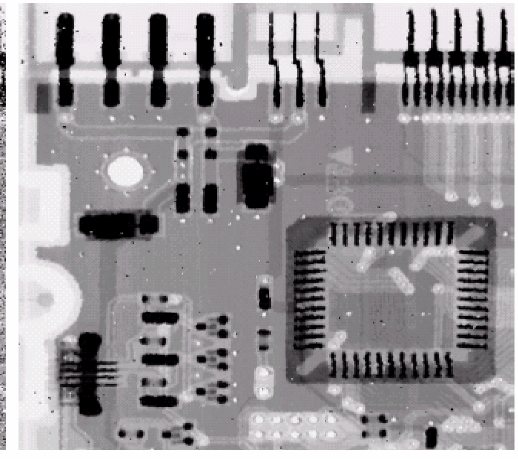
P_a=P_b=0.1的脉冲噪声

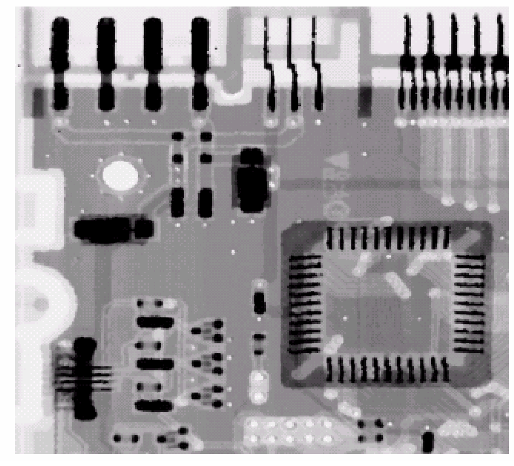
3×3的中值滤波器

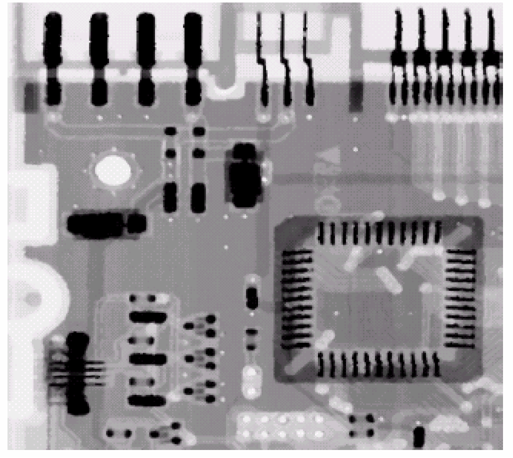
a b
c d

FIGURE 5.10
(a) Image corrupted by salt-and-pepper noise with probabilities $P_a = P_b = 0.1$.
(b) Result of one pass with a median filter of size 3 × 3.
(c) Result of processing (b) with this filter.
(d) Result of processing (c) with the same filter.









第二次中值滤波器处理

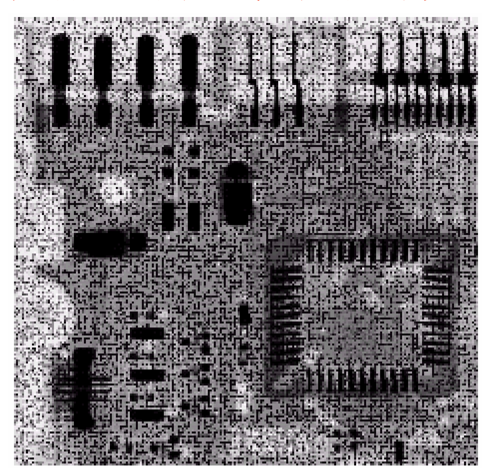
第三次中值滤波器处理，全部噪声消除

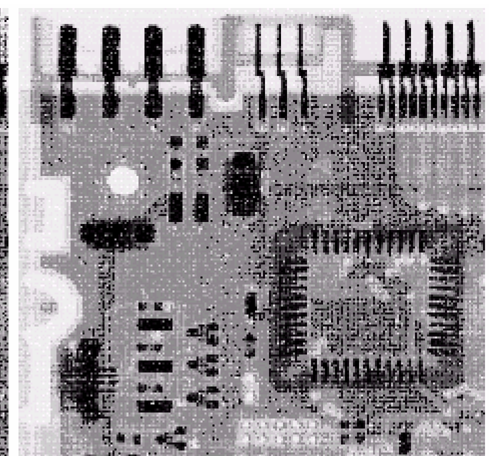
最大值和最小值滤波器举例

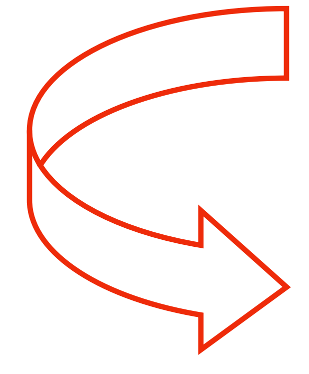


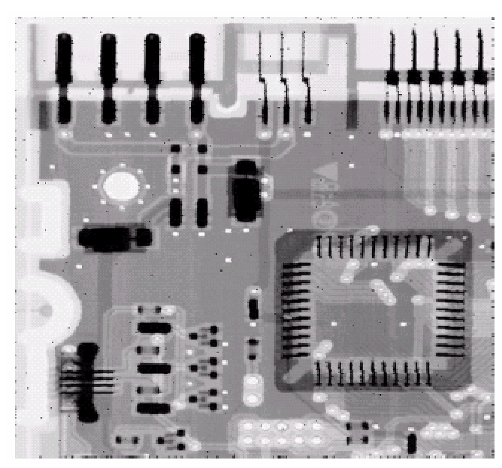
“胡椒”噪声干扰图像

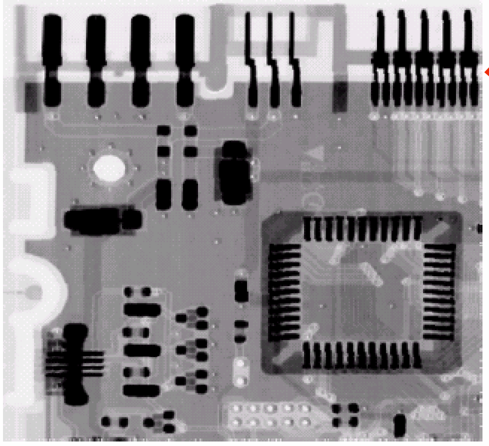












最大值滤波器处理　　最小值滤波器处理

a b

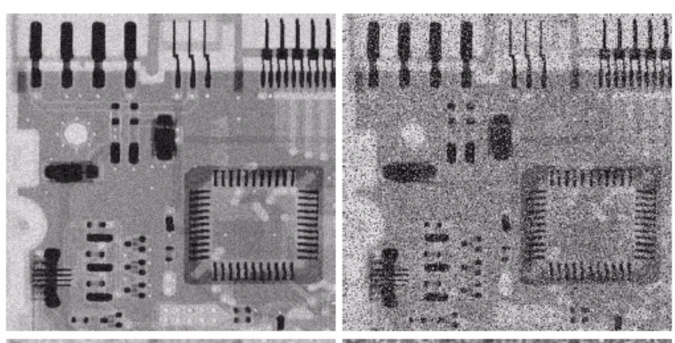
FIGURE 5.11 (a) Result of filtering Fig. 5.8(a) with a max filter of size 3 × 3. (b) Result of filtering 5.8(b) with a min filter of the same size.

空间域滤波器举例



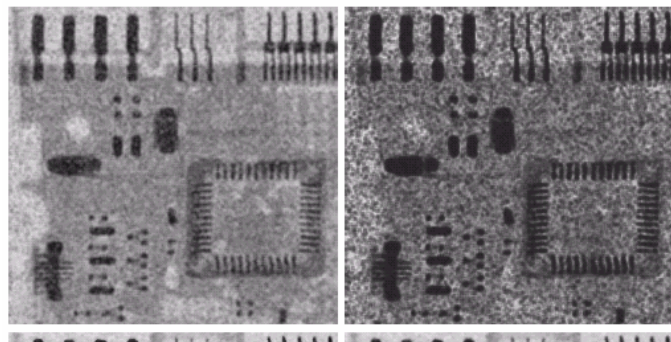
由于脉冲噪声的存在，算术均值和几何均值滤波器没有起到好的作用

均值为0，方差为800的噪声干扰的图像



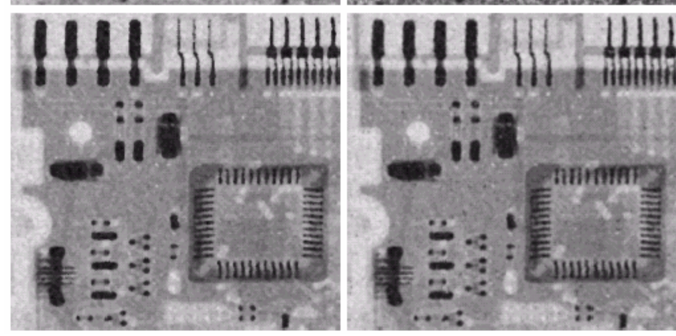
被$P_a=P_b=0.1$的椒盐噪声叠加，进一步恶化

算术均值滤波器 x



几何均值滤波器×

中值滤波器



d=5, 规格为5×5的修正后的阿尔法均值滤波器



自适应滤波器



- 自适应滤波器
  - ✓ 行为变化基于由m×n矩形窗口S_{xy}定义的区域内图像的统计特性
  - ✓ 与前述滤波器相比，性能更优
  - ✓ 但也增加了算法复杂性
  - ✓ 包括：
    - ➢ 自适应、局部噪声消除滤波器
    - ➢ 自适应中值滤波器

自适应滤波器



- 自适应、局部噪声消除滤波器

  滤波器响应基于以下4个量：

  - ✓ g(x, y), 表示噪声图像在点（x, y）上的值
  - ✓ σ_η^2 , g(x, y)的噪声方差
  - ✓ m_L， 在S_{xy}上像素点的局部均值
  - ✓ σ_L^2 , 在S_{xy}上像素点的局部方差

自适应滤波器



- 自适应、局部噪声消除滤波器

滤波器的预期性能如下：

1. 如果σ_η^2 ＝0，滤波器返回g(x, y)的值。因为在g(x, y)下零噪声的情况等同于f(x, y)
2. 如果局部方差σ_L^2与 σ_η^2 高相关，滤波器返回一个g(x, y)的近似值
3. 如果$\sigma_\eta^2 = \sigma_L^2$，滤波器返回区域$S_{xy}$上像素的算术均值。这样局部噪声用求平均$m_L$来降低

自适应滤波器



- 自适应、局部噪声消除滤波器

基于上述假定的自适应表达式：

$$\hat{f}(x,y)=g(x,y)-\frac{\sigma_\eta^2}{\sigma_L^2}\left[g(x,y)-m_L\right]$$

✓ 唯一需要知道或估计的未知量是噪声方差σ_η^2

✓ 其它参数可以从S_{xy}中的像素计算出来

自适应滤波器举例



均值为0，方差为1000的高斯噪声　　7×7的算术均值滤波器

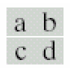
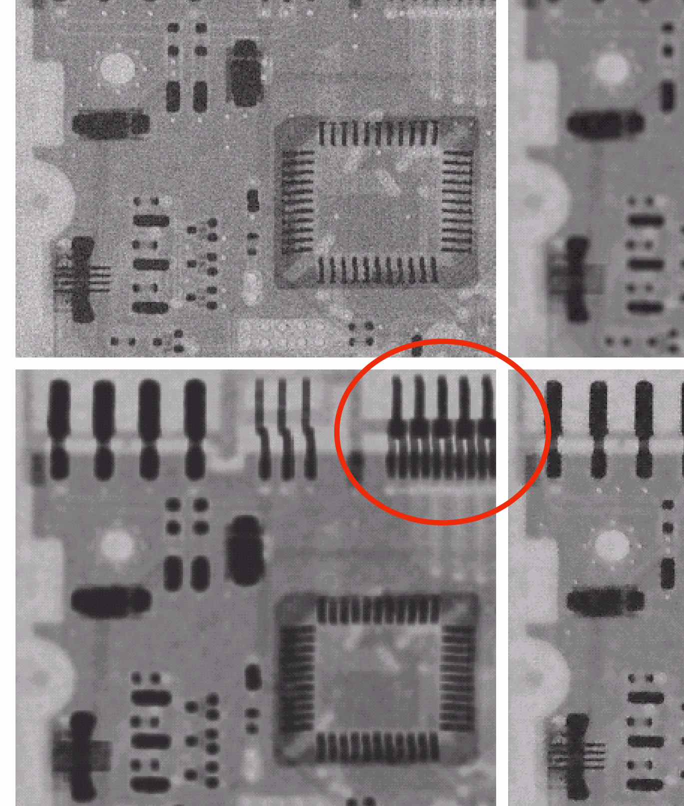
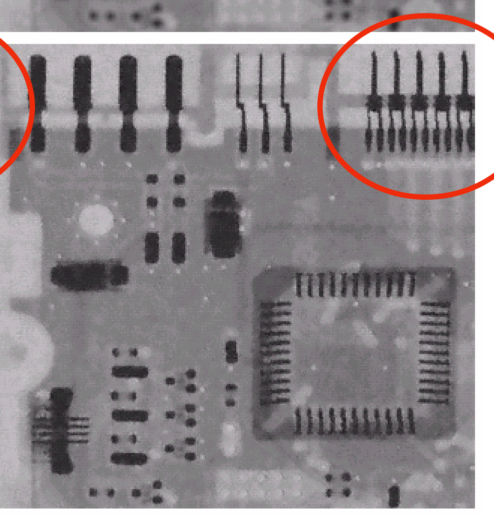


FIGURE 5.13 (a) Image corrupted by additive Gaussian noise of zero mean and variance 1000. (b) Result of arithmetic mean filtering. (c) Result of geometric mean filtering. (d) Result of adaptive noise reduction filtering. All filters were of size 7×7.





更加尖锐

7×7的几何均值滤波器　　7×7的自适应滤波器

自适应滤波器



- 自适应中值滤波器
  - ✓ 传统中值滤波器只能处理空间密度不大的冲激噪声（$p_a, p_b<0.2$），而自适应中值滤波器可以处理具有更大概率的冲激噪声
  - ✓ 可以在平滑非冲激噪声时保存细节，而传统中值滤波器无法做到

自适应滤波器



- 自适应中值滤波器：定义下列符号
  - ✓ z_{min}=S_{xy}中灰度级的最小值
  - ✓ z_{max}=S_{xy}中灰度级的最大值
  - ✓ z_{med}=S_{xy}中灰度级的中值
  - ✓ z_{xy}＝在坐标（x, y）上的灰度级
  - ✓ S_{max}=S_{xy}允许的最大尺寸

自适应中值滤波器：算法



- 主要目的
  - 除去“椒盐”噪声（冲激噪声）
  - 平滑其它非冲激噪声
  - 减少物体边界细化或粗化等失真

自适应中值滤波器：算法



- A层：找到一个非脉冲的中值
  - ✓ $A1=z_{med}-z_{min}$
  - ✓ $A2=z_{med}-z_{max}$
  - ✓ 如果A1>0且A2<0，转到B层，否则增大窗口尺寸
  - ✓ 如果窗口尺寸≤S_{max}，重复A层，否则输出z_{xy}

满足$z_{min}<z_{med}<z_{max}$，转到B层，说明z_{med}不是脉冲

- B层：
  - ✓ $B1=z_{xy}-z_{min}$
  - ✓ $B2=z_{xy}-z_{max}$
  - ✓ 如果B1>0且B2<0，输出z_{xy}，否则输出z_{med}

满足$z_{min}<z_{xy}<z_{max}$，说明z_{xy}不是脉冲，输出z_{xy}

说明$z_{xy}=z_{min}$或$z_{xy}=z_{max}$，输出z_{med}

自适应中值滤波器举例



$P_a=P_b=0.25$的椒盐噪声　　7×7的中值滤波器　　$S_{max}=7$的自适应中值滤波器

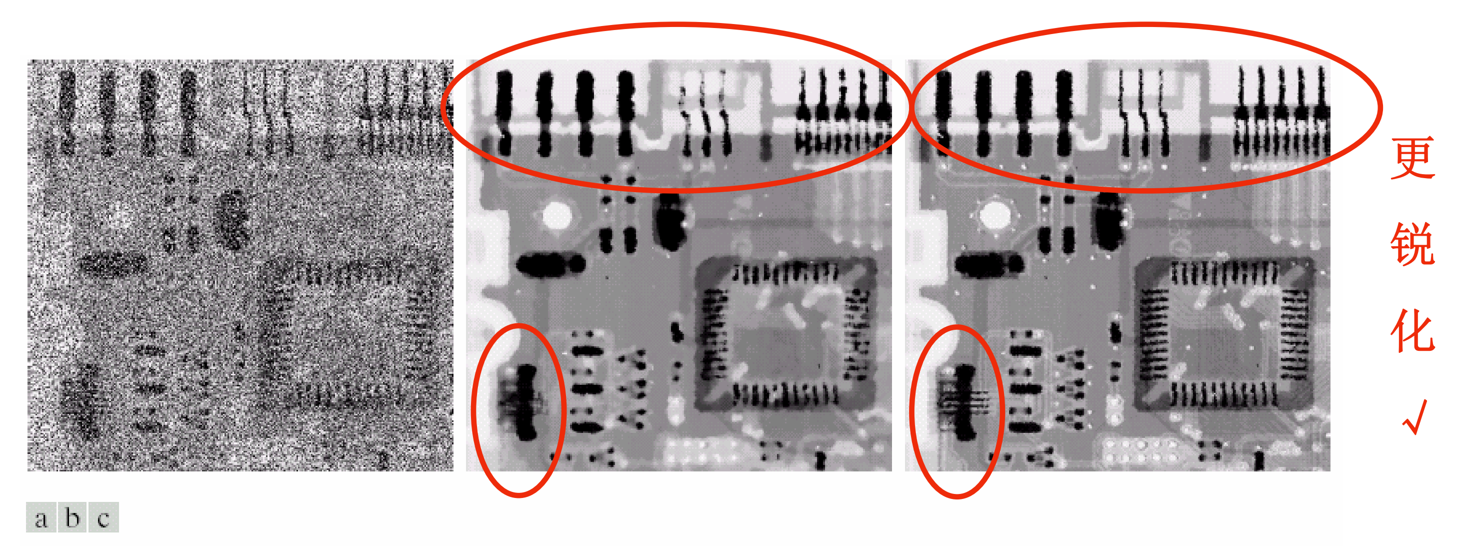


FIGURE 5.14 (a) Image corrupted by salt-and-pepper noise with probabilities $P_a = P_b = 0.25$. (b) Result of filtering with a 7 × 7 median filter. (c) Result of adaptive median filtering with $S_{max} = 7$.

更锐化√

图像复原



- 图像复原的频率域滤波器
  - ✓ 带阻滤波器
  - ✓ 带通滤波器
  - ✓ 陷波滤波器
  - ✓ 最佳陷波滤波器

图像复原



- 带阻滤波器：阻止一定频率范围内的信号通过而允许其它频率范围内的信号通过，消除或衰减傅里叶变换原点处的频段
  - ✓ 理想带阻滤波器
  - ✓ 巴特沃思带阻滤波器
  - ✓ 高斯带阻滤波器

图像复原



- 理想带阻滤波器

$$H(u,v)=\begin{cases}1, & D(u,v)< D_0-\frac{W}{2}\\ 0, & D_0-\frac{W}{2}\le D(u,v)\le D_0+\frac{W}{2}\\ 1, & D(u,v)> D_0+\frac{W}{2}\end{cases}$$

$$D(u,v)=\left[\left(u-\frac{M}{2}\right)^2+\left(v-\frac{N}{2}\right)^2\right]^{\frac{1}{2}}$$

W是频带的宽度，D_0是频带的中心半径

图像复原



- n阶的巴特沃思带阻滤波器

$$H(u,v)=\frac{1}{1+\left[\frac{D(u,v)W}{D^2(u,v)-D_0^2}\right]^{2n}}$$

- 高斯带阻滤波器

$$H(u,v)=1-e^{-\frac{1}{2}\left[\frac{D^2(u,v)-D_0^2}{D(u,v)W}\right]^2}$$

带阻滤波器的透视图



理想　　巴特沃思(阶数为1)　　高斯

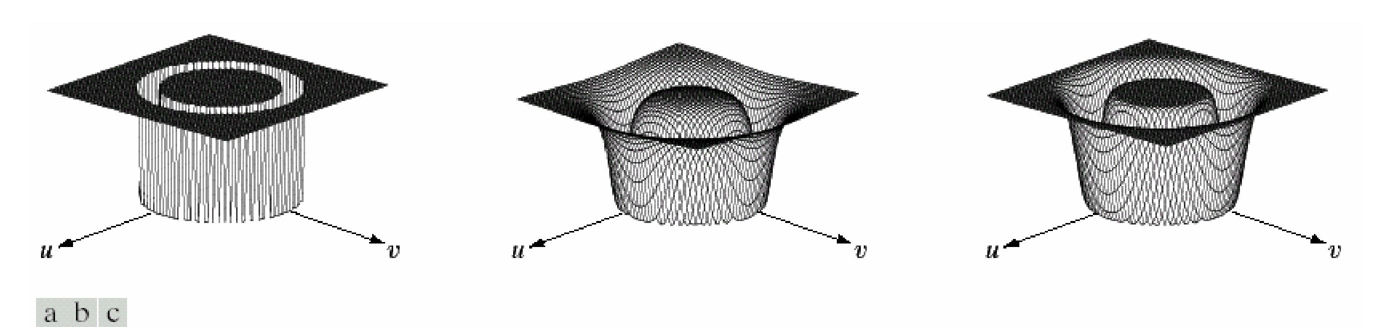
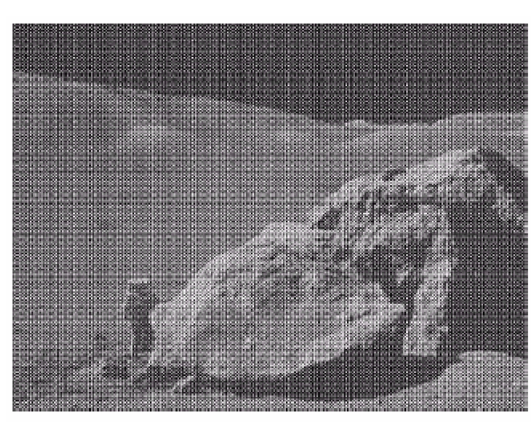


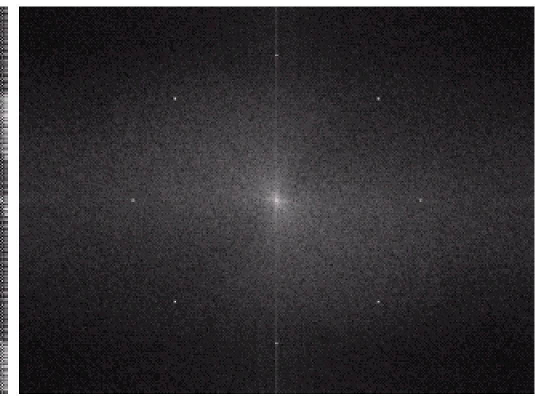
FIGURE 5.15 From left to right, perspective plots of ideal, Butterworth (of order 1), and Gaussian bandreject filters.

带阻滤波器使用举例



被正弦噪声污染的图像　　该图像的傅里叶频谱





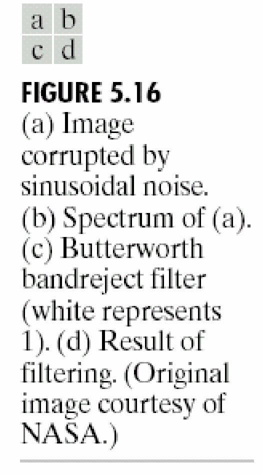
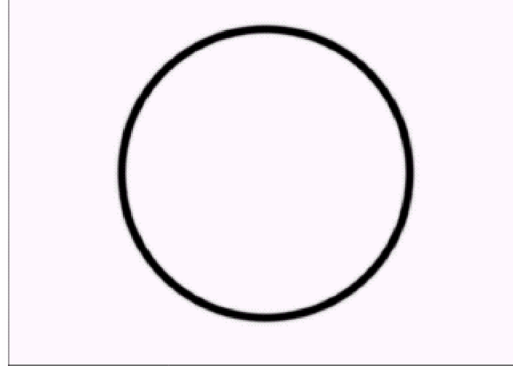


FIGURE 5.16 (a) Image corrupted by sinusoidal noise. (b) Spectrum of (a). (c) Butterworth bandreject filter (white represents 1). (d) Result of filtering. (Original image courtesy of NASA.)





4阶的巴特沃思带阻滤波器　　滤波后的图像

图像复原



- 带通滤波器：允许一定频率范围内的信号通过而阻止其它频率范围内的信号通过

$$H_{bp} = 1 - H_{br}(u,v)$$

$H_{bp}(u, v)$表示带通滤波器，$H_{br}(u, v)$表示相应的带阻滤波器

图像复原



- 陷波滤波器
  - ✓ 阻止或通过事先定义的中心频率邻域内的频率
  - ✓ 由于傅里叶变换是对称的，陷波滤波器必须以关于原点对称的形式出现
  - ✓ 如果陷波滤波器位于原点处，则以它本身形式出现

图像复原



理想的陷波滤波器

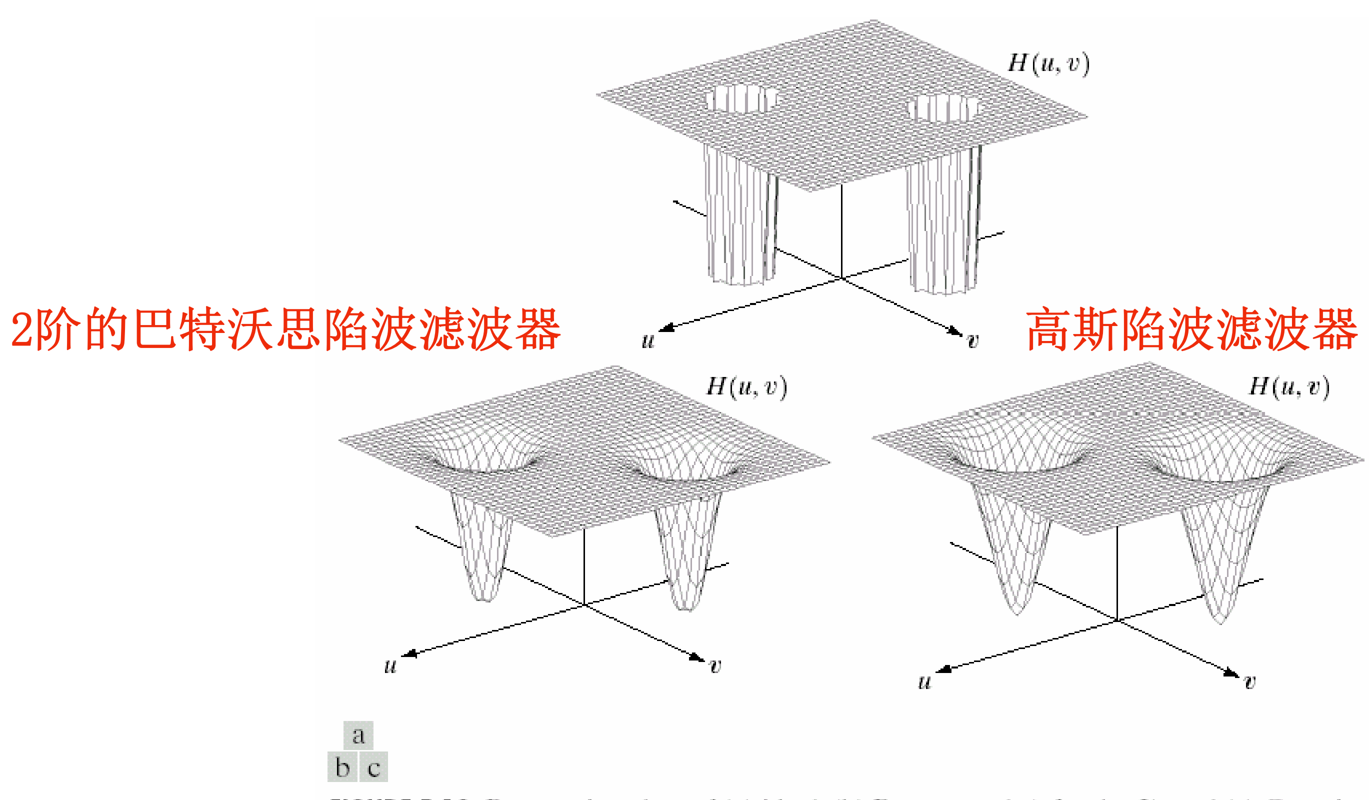


FIGURE 5.18 Perspective plots of (a) ideal, (b) Butterworth (of order 2), and (c) Gaussian notch (reject) filters.

图像复原



- 理想陷波带阻滤波器

$$H(u,v)=\begin{cases}0, D_1(u,v)\le D_0 或 D_2(u,v)\le D_0\\ 1 \quad 其它\end{cases}$$

$$D_1(u,v)=\left[\left(u-\frac{M}{2}-u_0\right)^2+\left(v-\frac{N}{2}-v_0\right)^2\right]^{\frac{1}{2}}$$

$$D_2(u,v)=\left[\left(u-\frac{M}{2}+u_0\right)^2+\left(v-\frac{N}{2}+v_0\right)^2\right]^{\frac{1}{2}}$$

中心在（u_0, v_0）且在（$-u_0, -v_0$）对称

图像复原



- 巴特沃思陷波带阻滤波器

$$H(u,v)=\frac{1}{1+\left[\frac{D_0^2}{D_1(u,v)D_2(u,v)}\right]^n}$$

- 高斯陷波带阻滤波器

$$H(u,v)=1-e^{-\frac{1}{2}\left[\frac{D_1(u,v)D_2(u,v)}{D_0^2}\right]}$$

注：当$u_0=v_0=0$, 上述3个滤波器变为高通滤波器

图像复原



- 陷波带通滤波器：通过包含在陷波区的频率

$$H_{np}(u,v)=1-H_{nr}(u,v)$$

✓ $H_{np}(u, v)$是陷波带通滤波器，$H_{nr}(u, v)$是对应的陷波带阻滤波器

✓ 当$u_0=v_0=0$时，陷波带通滤波器变为低通滤波器

数字图像处理（7）



任何问题？

研究生课程



数字图像处理

Digital Image Processing

彭宇新

北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

图像压缩



- 基本概念
- 图像压缩模型
- 信息论基础
- 无损压缩
- 有损压缩
- 图像压缩标准
- 视频压缩标准

图像压缩



- 基本概念
  - ✓ 概述
  - ✓ 数据冗余
    - ➢ 编码冗余
    - ➢ 像素间冗余
    - ➢ 心理视觉冗余
  - ✓ 图像保真度和质量

图像压缩



- 为什么需要图像压缩
  - ✓ 图像的数据量通常很大，对存储、处理和传输带来许多问题（对比视频）
  - ✓ 不断扩大的图像应用
    - ➢ Internet上的大量图像
    - ➢ 数字图书馆
    - ➢ 遥感图像
    - ➢ 视频，如电视会议、数字电视、IPTV
    - ➢ ……

图像压缩



- 图像压缩的方法
  - ✓ 消除冗余数据，从数学角度看，将原始图像转化为从统计角度看尽可能不相关的数据集
  - ✓ 一般分为两类：
    - ➢ 无损压缩：在压缩和解压缩过程中没有信息损失
    - ➢ 有损压缩：能取得较高的压缩率，但压缩后不能通过解压缩恢复原状
    - ➢ 其它：如根据需要，即可进行无损，也可进行有损压缩的技术；准无损技术

图像压缩



- 图像压缩的理论基础
  - ✓ 信息论
  - ✓ 图像处理的概念和技术
- 压缩方法
  - ✓ 预测编码方法（对应空域方法）
  - ✓ 变换编码方法（对应频域方法）

图像压缩



- 数据冗余的概念

数据是用来表示信息的。如果不同的方法为表示给定量的信息使用了不同的数据量，那么使用较多数据量的方法中，有些数据必然是代表了无用的信息，或者是重复地表示了其它数据已表示的信息，这就是数据冗余的概念。

图像压缩



- 相对数据冗余的定义（续）

如果n_1和n_2代表两个表示相同信息的数据集合中所携载信息单元的数量，则n_1表示的数据集合的相对数据冗余R_D定义为：

$$R_D = 1 - \frac{1}{C_R}$$

C_R称为压缩率，定义为

$$C_R = \frac{n_1}{n_2}$$

图像压缩



- 相对数据冗余和压缩率的一些特例

n1相对于n2	C_R	R_D	对应的情况
n1 = n2	1	0	第1种表达相对第2种表达不含冗余数据
n1 >> n2	$\to\infty$	$\to 1$	第1种数据集合包含相当多的冗余数据
n1 << n2	$\to 0$	$\to\infty$	第2种数据集合包含相当多的冗余数据

图像压缩



- 三种基本的数据冗余
  - ✓ 编码冗余
  - ✓ 像素间冗余
  - ✓ 心理视觉冗余
- 如果能减少或消除上述三种冗余的1种或多种冗余，就能取得数据压缩的效果

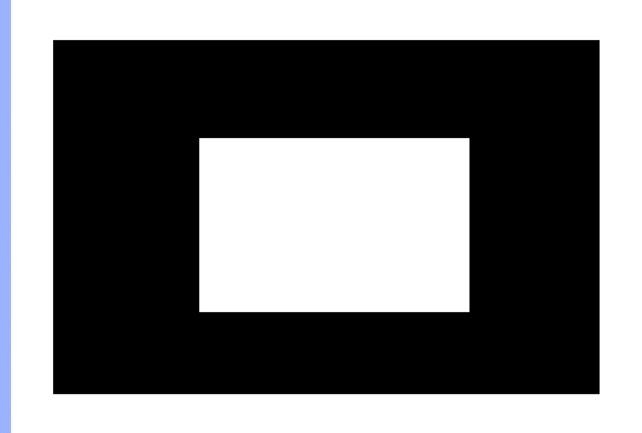
图像压缩



- 什么是编码冗余？

　　如果一个图像的灰度级编码，使用了多于实际需要的编码符号，就称该图像包含了编码冗余

- 黑白二值图像编码



如果用8位表示该图像的像素，我们就说该图像存在编码冗余，因为该图像的像素只有两个灰度，用一位即可表示。

编码冗余



- 图像直方图的定义

$$p_r(r_k)=\frac{n_k}{n} \qquad k=0,1,2,\ldots,L-1$$

n_k是第k个灰度级在图像中出现的次数，n是图像中的像素总数，L是灰度级数。

- 如果用于表示每个r_k值的比特数为$l(r_k)$，则表达每个像素所需的平均比特数为：

$$L_{avg}=\sum_{k=0}^{L-1}l(r_k)p_r(r_k)$$

表示灰度级所用的比特数和灰度级出现的概率相乘

编码冗余



- 对M×N的图像进行编码所需的比特数为

  MNL_{avg}

- 参考page 328的例8.1

像素间冗余



- 什么是像素间冗余？
  - ✓ 反映图像中像素之间的相互关系
  - ✓ 因为任何给定像素的值可以根据与这个像素相邻的像素进行预测，所以单个像素携带的信息相对较少
  - ✓ 对于一幅图像，很多单个像素对视觉的贡献是冗余的。它的值可以通过与它相邻的像素值为基础进行预测

例：原图像数据：234 223 231 238 235

压缩后数据：234 -11 8 7 -3

心理视觉冗余



- 什么是心理视觉冗余？

  ✓ 人眼感觉到的图像区域亮度不仅取决于该区域的反射光，例如根据马赫带效应，在灰度值为常数的区域也能感觉到灰度值的变化

  ✓ 这是由于眼睛对所有视觉信息感受的灵敏度不同。在正常视觉处理过程中各种信息的相对重要程度不同

  ✓ 有些信息在通常的视觉过程中与另外一些信息相比并不那么重要，这些信息被认为是心理视觉冗余的，去除这些信息并不会明显降低图像质量



心理视觉冗余



- 什么是心理视觉冗余？（续）

  ✓ 由于消除心理视觉冗余数据会导致一定量信息的丢失，所以这一过程通常称为量化

  ✓ 心理视觉冗余压缩是不可恢复的，量化的结果导致了数据有损压缩





保真度准则



- 保真度准则
  - ✓ 图像压缩可能会导致信息损失，如去除心理视觉冗余数据
  - ✓ 需要评价信息损失的测度以描述解码图像相对于原始图像的偏离程度，这些测度称为保真度准则
  - ✓ 常用保真度准则分为两大类：
    - ➢ 客观保真度准则
    - ➢ 主观保真度准则

保真度准则



- 客观保真度准则
  - ✓ 当所损失的信息量可以用编码输入图像与编码输出图像的函数表示时，它就是基于客观保真度准则的
  - ✓ 常用的两种客观保真度准则
    - ➢ 均方根误差
    - ➢ 均方信噪比

客观保真度准则



- 输入图和输出图之间的均方根误差

  ✓ 令 $f(x,y)$ 代表输入图，$\hat{f}(x,y)$ 代表对 $f(x,y)$ 先压缩后解压缩后得到的 $f(x,y)$ 的近似，则 $f(x,y)$ 和 $\hat{f}(x,y)$ 之间的误差定义为

$$e(x,y)=\hat{f}(x,y)-f(x,y)$$

如两幅图像尺寸均为M×N，则它们的总误差为

$$\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left|\hat{f}(x,y)-f(x,y)\right|$$

客观保真度准则



- 输入图和输出图之间的均方根误差
  - ✓ 这样$f(x,y)$和$\hat{f}(x,y)$之间的均方根误差为

$$e_{rms} = \left[\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left[\hat{f}(x,y) - f(x,y) \right]^2 \right]^{\frac{1}{2}}$$

心理视觉冗余



- 输出图的均方信噪比
  - ✓ 如果将 $\hat{f}(x,y)$ 看作原始图 $f(x,y)$ 和噪声信号 $e(x,y)$ 的和，那么输出图的均方信噪比 SNR_{ms} 为

$$SNR_{ms} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\hat{f}(x,y)^2 \Big/ \sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left[\hat{f}(x,y)-f(x,y)\right]^2$$

均方根信噪比 SNR_{rms} 为

$$SNR_{rms} = \sqrt{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\hat{f}(x,y)^2 \Big/ \sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left[\hat{f}(x,y)-f(x,y)\right]^2}$$



主观保真度准则

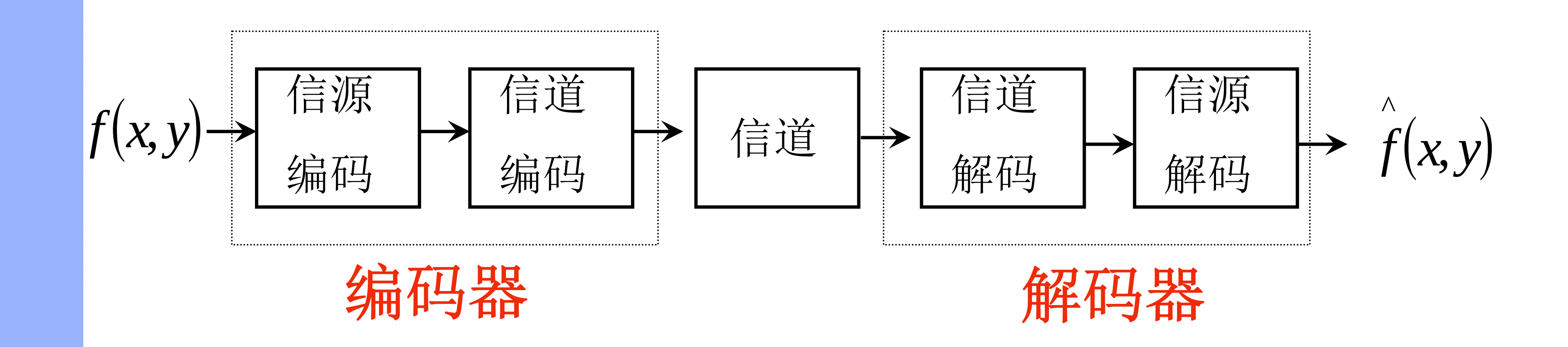


评分	评价	说明
1	优秀	图像质量非常好，如同人想象出的最好质量
2	良好	图像质量高，观看舒服，有干扰但不影响观看
3	可用	图像质量可接受，有干扰但不太影响观看
4	刚可看	图像质量差，干扰有些妨碍观看，希望改进
5	差	图像质量很差，妨碍观看的干扰始终存在，几乎无法观看
6	不能用	图像质量极差，不能使用

图像压缩模型



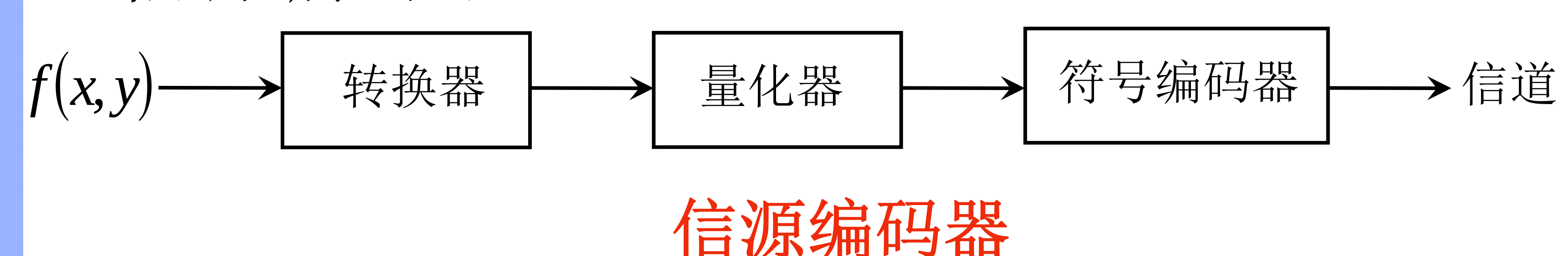
- 图像压缩模型



图像压缩模型



- 信源编码器



✓ 信源编码器：减少或消除输入图像中的编码冗余、像素间冗余及心理视觉冗余

✓ 转换器：减少像素间冗余

✓ 量化器：减少心理视觉冗余，该步操作是不可逆的

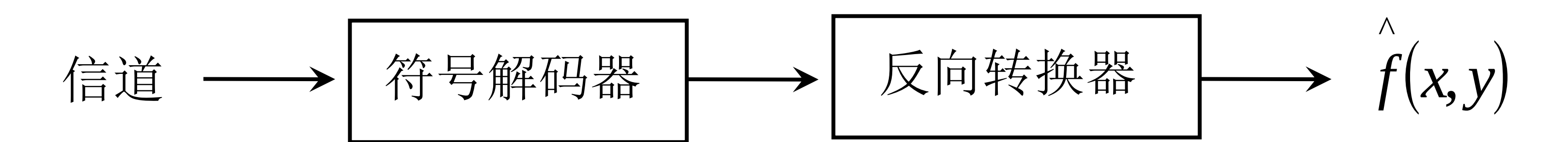
✓ 符号编码器：减少编码冗余

✓ 并不是每个图像压缩系统都必须包含这3种操作，如进行无误差压缩时，必须去掉量化器

图像压缩模型



- 信源解码器



信源解码器

✓ 符号解码器：进行符号编码的逆操作

✓ 反向转换器：进行转换器的逆操作

✓ 为什么没有反向量化器？

图像压缩模型



- 信道编码器和信道解码器

✓ 信道是有噪声的或易产生误差时，信道编码器和信道解码器对整个编解码过程非常重要

✓ 由于信源编码器的输出数据一般只有很少的冗余，所以它们对输出噪声很敏感

✓ 汉明(Hamming)编码：在编了码的码字后面增加足够的比特位以保证各个正确的码字之间至少有一定数量的比特位不相同

图像压缩模型



- 信息论基础
  - ✓ 显示一幅图像需要多大的数据量？有没有描述一幅图像且没有信息丢失的最小数据量？
- 信息测量
  - ✓ 对一个随机事件E，如果它的出现概率是P(E),那么它包含的信息：

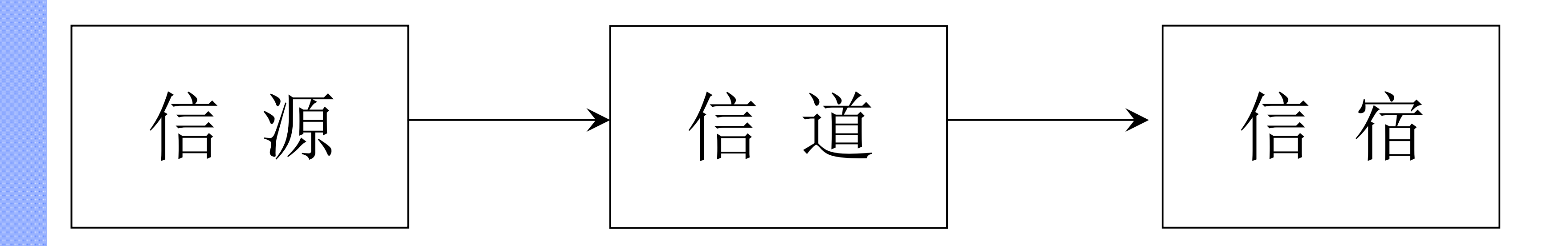
$$I(E)=\log\frac{1}{P(E)}=-\log P(E)$$

I(E)称为E的自信息。如果P(E)=1(即事件总发生)，那么I(E)=0

图像压缩模型



- 信息信道



Ensemble $(A, \mathbf{z})$

$A = \{a_j\}$

$\mathbf{z} = [P(a_1), P(a_2), \ldots, P(a_J)]^T$

$Q = [q_{kj}]$

Ensemble $(B, \mathbf{v})$

$B = \{b_k\}$

$\mathbf{v} = [P(b_1), P(b_2), \ldots, P(b_K)]^T$

信道是连接信源和用户的物理媒介。它可以是电话线、无线传播、导线或internet

信息论基础



- 信源
  - ✓ $A=\{a_1, a_2, \ldots, a_j\}$称为信源字母表
  - ✓ 信源产生符号a_j的事件概率是$P(a_j)$，且

$$\sum_{j=1}^{J} P(a_j) = 1$$

  - ✓ 一个J×1向量 $z=[P(a_1),P(a_2),\ldots,P(a_J)]^T$ 用于表示所有信源符号的概率集合
  - ✓ 有限总体集合(A, z)完全描述了信息源

信息论基础



- 信源（续）

  ✓ 如果产生k个信源符号，则大数定律保证对于一个充分大的k，符号a_j将被输出$kP(a_j)$次。因此，根据k输出得到的平均自信息是

$$-k\sum_{j=1}^{J}P(a_j)\log P(a_j)$$

  ✓ 每个信源输出的平均信息，也称为信源的熵为

$$H(z)=-\sum_{j=1}^{J}p(a_j)\log P(a_j)$$

  ✓ 如果信源符号的出现是等可能性的，则上述熵被最大化，此时信源提供最大信息量

信息论基础



- 信道输出
  - ✓ $B=\{b_1, b_2, \ldots, b_k\}$称为信道字母表
  - ✓ 提交给用户的字符b_k的概率是$P(b_k)$
  - ✓ 有限集合(B, v)完整描述了信道输出和用户收到的信息，$v=[P(b_1),P(b_2),\ldots,P(b_K)]^T$
  - ✓ 给定信道输出概率$P(b_K)$和信源符号概率$P(a_J)$，它们由下式相联系

$$P(b_k)=\sum_{j=1}^{J} P(b_k \mid a_j)P(a_j)$$

信息论基础



- 信道输出（续）

  ✓ 将上式中的条件概率放入一个K×J的正向信道传递矩阵Q，其元素$q_{kj}=P(b_k|a_j)$为条件概率

$$Q=\begin{bmatrix} P(b_1|a_1) & P(b_1|a_2) & \dots & P(b_1|a_J) \\ P(b_2|a_1) & P(b_2|a_2) & \dots & P(b_2|a_J) \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ P(b_K|a_1) & P(b_K|a_2) & \dots & P(b_K|a_J) \end{bmatrix}$$

  ✓ 则输出符号集的概率分布由下式计算

$$v = Qz$$

信息论基础



- 条件熵函数H(z|b_k)

$$H(z\,|\,b_k)=-\sum_{j=1}^{J}P(a_j\,|\,b_k)\log P(a_j\,|\,b_k)$$

- 条件概率H(z|v)

$$H(z\,|\,v)=\sum_{k=1}^{K}H(z\,|\,b_k)P(b_k)$$

$$=-\sum_{j=1}^{J}\sum_{k=1}^{K}P(a_j\,|\,b_k)\log P(a_j\,|\,b_k)P(b_k)$$

$$=-\sum_{j=1}^{J}\sum_{k=1}^{K}P(a_j\,|\,b_k)P(b_k)\log P(a_j\,|\,b_k)$$

$$=-\sum_{j=1}^{J}\sum_{k=1}^{K}P(a_j,b_k)\log P(a_j\,|\,b_k)$$

信息论基础



- 因为两个事件C和D的联合概率是

$$P(C,D)=P(C|D)P(D)=P(D|C)P(C)$$

- $P(a_j)$的变换（下面推导互信息使用）

$$P(a_j)=P(a_j,b_1)+P(a_j,b_2)+\ldots+P(a_j,b_K)$$

$$=\sum_{k=1}^{K}P(a_j,b_K)$$

- 定义信道传输元素

$$q_{kj}=P(b_k|a_j)$$

信息论基础



- z和v的互信息定义为

$$I(z,v)=H(z)-H(z|v)$$

$$=\left(-\sum_{j=1}^{J}P(a_j)\log P(a_j)\right)-\left(-\sum_{j=1}^{J}\sum_{k=1}^{K}P(a_j,b_k)\log P(a_j|b_k)\right)$$

$$=\sum_{j=1}^{J}\sum_{k=1}^{K}P(a_j,b_k)\log P(a_j|b_k)-\sum_{j=1}^{J}\sum_{k=1}^{K}P(a_j,b_k)\log P(a_j)$$

$$=\sum_{j=1}^{J}\sum_{k=1}^{K}P(a_j,b_k)\log\frac{P(a_j|b_k)}{P(a_j)}$$

$$=\sum_{j=1}^{J}\sum_{k=1}^{K}P(a_j,b_k)\log\frac{P(a_j|b_k)P(b_k)}{P(a_j)P(b_k)}$$

$$=\sum_{j=1}^{J}\sum_{k=1}^{K}P(a_j,b_k)\log\frac{P(a_j,b_k)}{P(a_j)P(b_k)}$$





- z和v的互信息的另外一种表达

$$
\begin{aligned}
I(z,v) &= \sum_{j=1}^{J}\sum_{k=1}^{K} P(a_j,b_k)\log\frac{P(a_j,b_k)}{P(a_j)P(b_k)} \\
&= \sum_{j=1}^{J}\sum_{k=1}^{K} P(b_k \mid a_j)P(a_j)\log\frac{P(b_k \mid a_j)P(a_j)}{P(a_j)P(b_k)} \\
&= \sum_{j=1}^{J}\sum_{k=1}^{K} P(a_j)q_{kj}\log\frac{q_{kj}}{P(b_k)} \\
&= \sum_{j=1}^{J}\sum_{k=1}^{K} P(a_j)q_{kj}\log\frac{q_{kj}}{\sum_{i=1}^{J} P(b_k \mid a_i)P(a_i)} \\
&= \sum_{j=1}^{J}\sum_{k=1}^{K} P(a_j)q_{kj}\log\frac{q_{kj}}{\sum_{i=1}^{J} P(a_i)q_{ki}}
\end{aligned}
$$

信息论基础



- 互信息总结
  - ✓ 互信息I(z, v)是信源符号概率向量z和信道矩阵Q的函数
  - ✓ 当输入和输出符号统计独立时，I(z, v)取得最小值0
  - ✓ I(z, v)对所有信源分布u的最大值就是信道容量 $C=\max_{u}\left[I(u,v)\right]$
  - ✓ 信道容量定义了能够通过信道可靠地传送信息的最大传送率

信息论基础



- 基本编码定理（自学）
  - ✓ 无噪声编码定理
  - ✓ 噪声编码定理
  - ✓ 信源编码定理

信息论基础



- 无误差压缩
  - ✓ 变长编码
    - ➢ 霍夫曼(Huffman)编码
    - ➢ 其它变长编码
    - ➢ 算术编码
  - ✓ LZW编码
  - ✓ 位平面编码
  - ✓ 无损预测编码

无误差压缩



- 无误差压缩的必要性
  - ✓ 在医疗或商业文件的归档，有损压缩因为法律原因而被禁止
  - ✓ 卫星成像的收集，考虑数据使用和所花费用，不希望有任何数据损失
  - ✓ X光拍片，信息的丢失会导致诊断的正确性
  - ✓ ……
- 无误差压缩技术
  - ✓ 减少像素间冗余
  - ✓ 减少编码冗余

无误差压缩



- 变长编码
  - ✓ 减少编码冗余
  - ✓ 变长编码，即把最短的码字赋予出现概率最大的灰度级
- 霍夫曼编码
  1. 将需要考虑的符号概率排序，并将最低概率的符号联结为一个单一符号
  2. 对每个化简后的信源进行编码，从最小的信源开始，一直编码到原始的信源

霍夫曼编码



步骤1

Original source		Source reduction			
Symbol	Probability	1	2	3	4
a_2	0.4	0.4	0.4	0.4	0.6
a_6	0.3	0.3	0.3	0.3	0.4
a_1	0.1	0.1	0.2	0.3	
a_4	0.1	0.1	0.1		
a_3	0.06	0.1			
a_5	0.04				

FIGURE 8.11 Huffman source reductions.

步骤2

FIGURE 8.12 Huffman code assignment procedure.

Original source			Source reduction							
Sym.	Prob.	Code	1		2		3		4	
a_2	0.4	1	0.4	1	0.4	1	0.4	1	0.6	0
a_6	0.3	00	0.3	00	0.3	00	0.3	00	0.4	1
a_1	0.1	011	0.1	011	0.2	010	0.3	01		
a_4	0.1	0100	0.1	0100	0.1	011				
a_3	0.06	01010	0.1	0101						
a_5	0.04	01011								

霍夫曼编码



- 霍夫曼解码
  - 解码通过查询表的方式完成
  - 例：编码串01010011 1 1 100

a_3 a_1 $a_2a_2a_6$

霍夫曼编码



- 其它接近最佳的变长编码：为什么需要？
  - ✓ 当对大量符号进行编码，构造霍夫曼编码比较复杂
  - ✓ 对J个信源符号，需要进行J-2次信源化简和J-2次编码分配
  - ✓ 对256个灰度级图像，需要254次信源化简和254次编码分配
  - ✓ 考虑牺牲编码效率以减少编码构造的复杂性

算术编码



- 算术编码
  - ✓ 从整个符号序列出发，采用递推形式连续编码
  - ✓ 在算术编码中，源符号和码字间的一一对应关系并不存在。1个算术码字要赋给整个信源符号序列，而码字本身确定0和1之间的1个实数区间
  - ✓ 随着符号序列中的符号数量增加，用来代表它的区间减小而表达区间的信息单位数量变大



算术编码



Source Symbol	Probability	Initial Subinterval
a_1	0.2	[0.0, 0.2)
a_2	0.2	[0.2, 0.4)
a_3	0.4	[0.4, 0.8)
a_4	0.2	[0.8, 1.0)

TABLE 8.6 Arithmetic coding example.

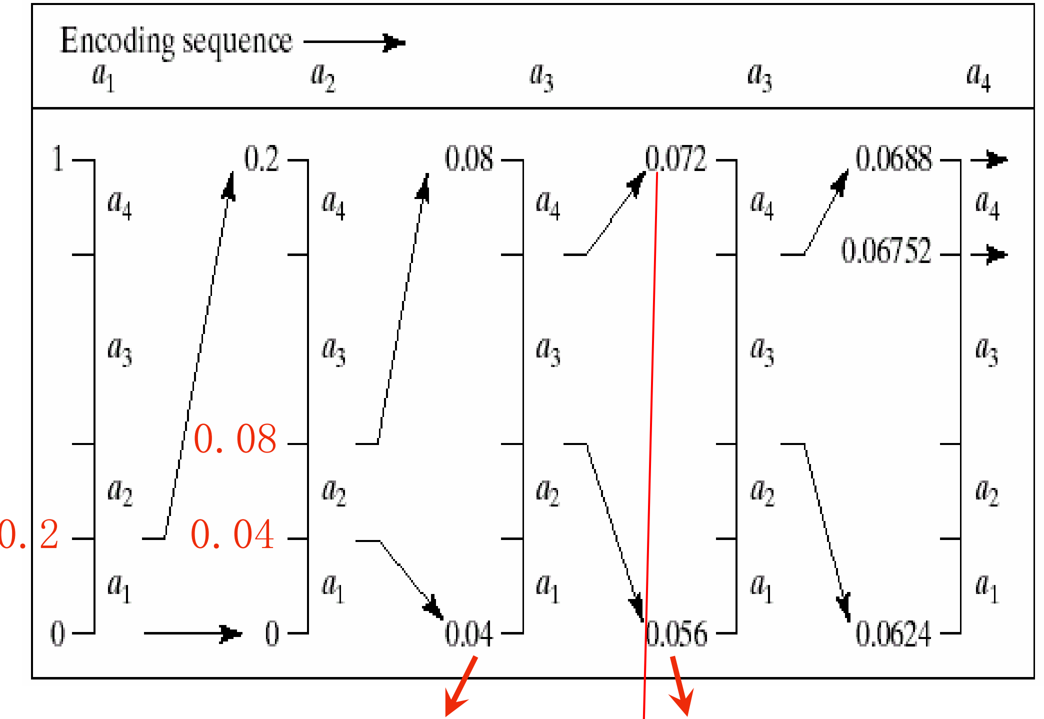


FIGURE 8.13 Arithmetic coding procedure.

$0.04=0+(0.2-0)/5\times1$

$0.08=0+(0.2-0)/5\times2$

$0.056=0.04+(0.08-0.04)/5\times2$

$0.072=0.04+(0.08-0.04)/5\times4=0.04+0.032$

LZW编码



- LZW编码：消除像素间冗余
  - ✓ 是由Lemple和Ziv最早提出，然后由Welch充实的有专利保护的LZW算法
  - ✓ 将原始数据中的重复字符串建立一个字串表，然后用该重复字串在字串表中的索引替代原始数据达到压缩的目的
  - ✓ 一个由8位组成的文件可以被编成12位的代码。在2^{12}＝4096个可能的代码中，2^8＝256个代表所由可能的单个字符（8位），剩下的3840个代码分配给压缩过程中出现的字符串

LZW编码



- LZW编码（续）
  - 每当表中没有的字符串第一次出现时，它就被原样保存，同时将分配给它的代码也保存
  - 之后，当这个串再次出现时，只将它的代码保存，这就去掉了文件冗余信息
  - 不但字符串表是在压缩过程中动态生成，而且字符串表也不必保存在压缩文件里，因为解压缩算法可以由压缩文件中的信息重构
  - 使用LZW的文件格式包括GIF，TIFF和PDF等。

LZW编码例子



- 一个4×4、8位图像

39	39	126	126
39	39	126	126
39	39	126	126
39	39	126	126

- 一个512字节的字典

字典位置	条　目
0	0
1	1
…	…
255	255
256	——
…	…
511	——

LZW编码例子



Currently Recognized Sequence	Pixel Being Processed	Encoded Output	Dictionary Location (Code Word)	Dictionary Entry
	39			
39	39	39	256	39-39
39	126	39	257	39-126
126	126	126	258	126-126
126	39	126	259	126-39
39	39			
39-39	126	256	260	39-39-126
126	126			
126-126	39	258	261	126-126-39
39	39			
39-39	126			
39-39-126	126	260	262	39-39-126-126
126	39			
126-39	39	259	263	126-39-39
39	126			
39-126	126	257	264	39-126-126
126		126		

Appear

39−39

126−126

39−39

39−39−126

126−39

39−126

TABLE 8.7 LZW coding example.

位平面编码



- 位平面编码：消除像素间冗余
  - ✓ 将一幅图像分解为一系列二值图像并通过二值图像压缩方法对每幅二值图像进行压缩
  - ✓ 位平面分解的两种方法
    - 二值图像位平面
    - 灰度编码位平面

位平面编码



- 二值图像位平面
  - 一幅m比特的灰度图像具有的灰度级表示如下

$$a_{m-1}2^{m-1} + a_{m-2}2^{m-2} + \ldots + a_1 2^1 + a_0 2^0$$

  - 零级位平面是通过收集每个像素的a_0位生成，第(m-1)级位平面包含a_{m-1}位
  - 缺点：图像在灰度级上稍有变化就会对位平面的复杂性产生显著影响，如亮度127(01111111)和亮度128(10000000)的转换

位平面编码



- 灰度编码位平面
  - ✓ 图像的灰度编码根据下列方法得到：

$$g_{m-1} = a_{m-1}$$

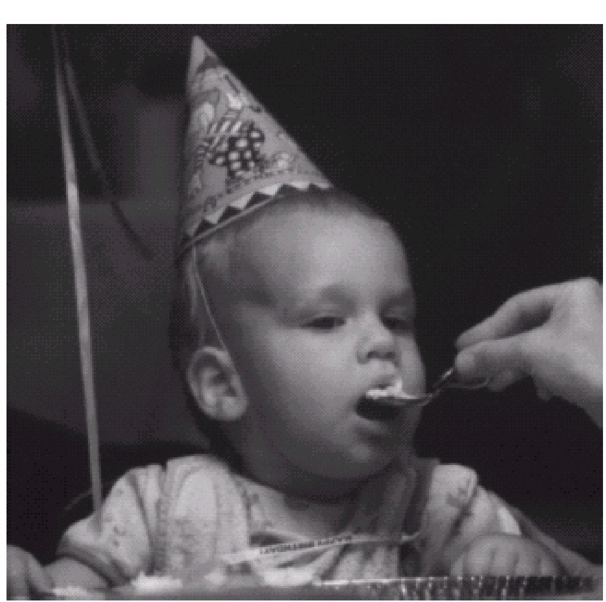
$$g_i = a_i \oplus a_{i+1} \qquad 0 \le i \le m-2$$

  - ✓ 避免二值图像位平面的问题，连续码字只在1位位置上不同，如亮度127 (01000000)和亮度128(11000000)的转换



位平面编码——例

8比特单色图像　　　　二值图像



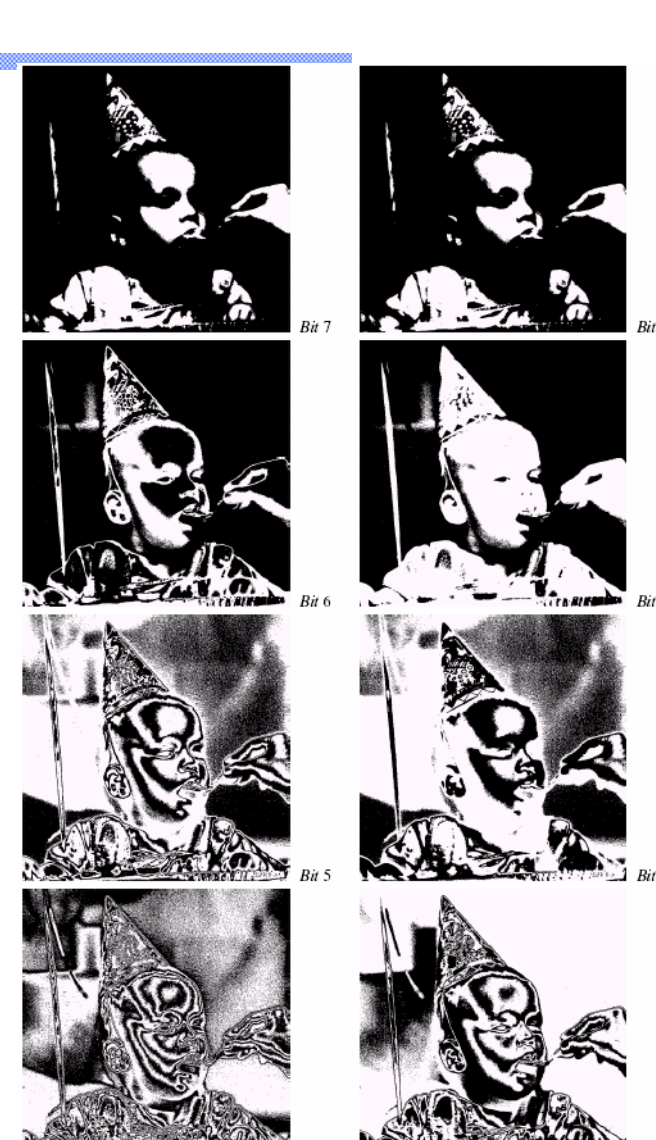
a b

FIGURE 8.14 A 1024 × 1024 (a) 8-bit monochrome image and (b) binary image.

位平面编码——例



二值图像位平面

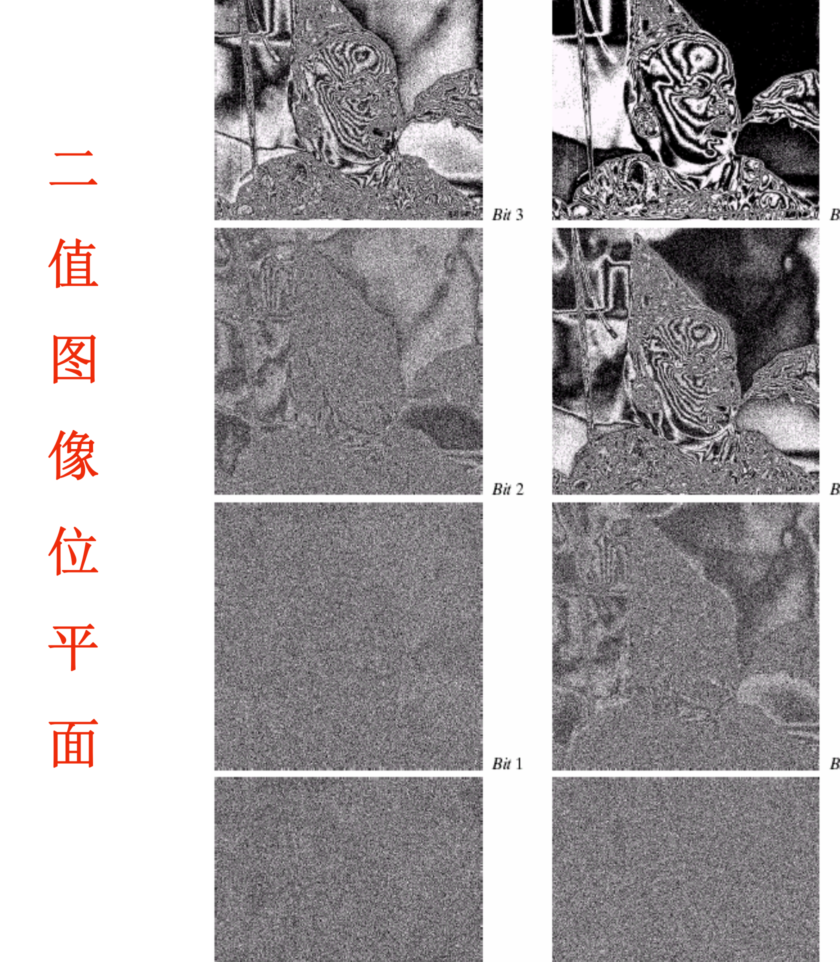


灰度编码位平面

位平面编码——例



二值图像位平面



灰度编码位平面

位平面编码



- 位平面分解方法总结
  - ✓ 低位面图比高位面图复杂，即低位面图比高位面图包括的细节要多，也更随机
  - ✓ 灰度编码表达的位面图复杂度较低，但具有视觉意义信息的位面图数量更多

位平面编码



- 常数块编码
  - ✓ 用专门的码字表达全是0或1的连通区域
  - ✓ 将图像分成全黑，全白或混合的m×n尺寸的块。出现频率最高的类赋予1位码字0，其它2类分别赋予2位码字10和11
  - ✓ 由于原来需用mn比特表示的常数块现在只用1位或2位表示，这样就达到了压缩的目的
  - ✓ 赋予混合块的码只是作为前缀，后面还需跟上该块的用mn位表示的模式

位平面编码



- 常数块编码（续）
  - ✓ 当需压缩的图像由白色部分组成时（如文档），可将白色区域编为0，其它块用1接上该块的位模式编码
  - ✓ 改进办法：将二值图或位平面迭代地分解成尺寸越来越小的子块。如果子块不是全白，继续分解，直至某个事先确定的子块尺寸。如果最后子块全白，就编为0，反之编为1加上该块的位模式

位平面编码



- 一维行程编码
  - ✓ 用一长度序列表示图像或位平面的每一行，这些长度描绘了对黑色和白色像素的连续行程，这称为行程编码
  - ✓ 是传真编码的标准压缩方法
  - ✓ 对从左到右扫描一行时所遇到的1或0的连接组，使用这些连接组的长度进行编码
  - ✓ 决定行程长度值的常用方法：
    - ➢ 指定每一行第一次行程的值
    - ➢ 假设每一行从白色行程开始，这次行程的长度可能为0

无损预测编码



- 预测编码的基本思想
  - ✓ 通过仅提取每个像素中的新信息并对它们编码来消除像素间的冗余
  - ✓ 1个像素的新信息定义为该像素的当前值与预测值的差
  - ✓ 正是由于像素间有相关性，所以才使预测成为可能





- 无损预测编解码系统

a
b

FIGURE 8.19 A lossless predictive coding model: (a) encoder; (b) decoder.

Input image
f_n
Predictor
Nearest integer
$\hat{f}_n$
$+\Sigma-$
e_n
Symbol encoder
Compressed image

Compressed image
Symbol decoder
e_n
$+\Sigma+$
f_n
Decompressed image
Predictor
$\hat{f}_n$

无损预测编码



- 当输入图像的像素序列f_n逐个进入编码器，预测器根据过去的输入产生当前输入像素的估计值。预测器的输出舍入成最近的整数$\hat{f}_n$并被用来计算预测误差

$$e_n = f_n - \hat{f}_n$$

- 该误差用符号编码器借助变长码进行编码以产生压缩数据流的下一个元素。然后解码器根据接收到的变长码字重建e_n，并执行下列操作

$$f_n = e_n + \hat{f}_n$$

无损预测编码



- 一般情况下，可将m个先前的像素进行线性组合以得到预测

$$\hat{f}_n = round\left[\sum_{i=1}^{m} a_i f_{n-i}\right]$$

$$\hat{f}_n(x, y) = round\left[\sum_{i=1}^{m} a_i f(x, y-i)\right]$$

其中，m是线性预测器的阶，round是舍入函数，a_i是预测系数

- 最简单的一维线性预测编码是一阶（m=1），此时

$$\hat{f}_n = round\left[af(x, y-1)\right]$$

数字图像处理（8）



任何问题？

研究生课程



数字图像处理

Digital Image Processing

彭宇新

北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

图像压缩



- 基本概念
- 图像压缩模型
- 信息论基础
- 无损压缩
- 有损压缩
- 图像压缩标准
- 视频压缩标准

有损压缩



- 概述
  - ✓ 牺牲图像复原的准确度以换取压缩能力的增加
  - ✓ 如果产生的失真可以容忍，则压缩能力的增加是有效的

有损压缩



- 有损预测编码：直接对像素在图像空间进行操作, 称为空域方法
  - ✓ 有损预测编码系统
  - ✓ 最优预测器
  - ✓ 最优量化
- 变换编码：基于图像变换的编码方法，称为频域方法
  - ✓ 变换编码系统
  - ✓ 变换选择
  - ✓ 子图像尺寸选择
  - ✓ 比特分配

有损预测编码系统



Input image → f_n → $+\Sigma-$ → e_n → Quantizer → $\dot{e}_n$ → Symbol encoder → Compressed image

Predictor ← $\dot{f}_n$ ← Σ (+ $\dot{e}_n$, + $\hat{f}_n$); Predictor → $\hat{f}_n$

Compressed image → Symbol decoder → $\dot{e}_n$ → $+\Sigma-$ → $\dot{f}_n$ → Decompressed image

Predictor → $\hat{f}_n$

a
b

FIGURE 8.21 A lossy predictive coding model: (a) encoder and (b) decoder.

$$\dot{f}_n = \dot{e}_n + \hat{f}_n$$

有损压缩



- 有损预测编码系统
  - ✓ 量化器插在符号编码器和预测误差产生处之间，把原来无损编码器中的整数舍入模块吸收了进来
  - ✓ 量化器将预测误差映射进输出$\dot{e}_n$中，$\dot{e}_n$确定了有损预测编码中的压缩量和失真量
  - ✓ 反馈环的输入是过去预测和与其对应的量化误差的函数

$$\dot{f}_n = \dot{e}_n + \hat{f}_n$$

有损压缩



- 最优预测器

✓ 在绝大多数预测编码中用到的最优预测器在满足限制条件

$$\dot{f}_n = \dot{e}_n + \hat{f}_n \approx e_n + \hat{f}_n = f_n$$

$$\hat{f}_n = \sum_{i=1}^{m} \alpha_i f_{n-1}$$

的情况下能最小化编码器的均方预测误差

$$E\{e_n^2\} = E\left\{\left[f_n - \hat{f}_n\right]^2\right\}$$

有损压缩



- 最优预测器（续）
  - 最优准则是最小化均方预测误差，设量化误差可以忽略（$\dot{e}_n \approx e_n$），并用m个先前像素的线性组合进行预测
  - 上述限制并不是必需的，但它们都极大地简化了分析，也减少了预测器的计算复杂性
  - 基于上述条件的预测编码方法称为差值脉冲码调制法（DPCM）
  - 最优预测器设计的问题简化为比较直观地选择m个预测系数以最小化下式的问题：

有损压缩



- 最优预测器（续）

$$E\{e_n^2\} = E\left\{\left[f_n - \sum_{i=1}^{m} \alpha_i f_{n-i}\right]^2\right\}$$

有损压缩



- 最优预测器（续）
  - ✓ 并假设用1个4阶线性预测器：

$$\hat{f}(x,y)=\alpha_1 f(x,y-1)+\alpha_2 f(x-1,y-1)+\alpha_3 f(x-1,y)+\alpha_4 f(x+1,y-1)$$

来预测

  - ✓ 通常，预测系数的和要小于或等于1。即

$$\sum_{i=1}^{m}\alpha_i \le 1$$

有损压缩



- 最优预测器（续）

$$\sum_{i=1}^{m}\alpha_i \le 1$$

✓ 上述限制是确保预测器的输出能够落到灰度级的允许范围内

✓ 减少传输噪声的影响

✓ 减小DPCM解码器对输入噪声的敏感性是很重要的，因为单个差错会传播到所有以后的输出。这样，解码器的输出会变得不稳定

例：预测技术的对比



- 考虑对单色图像进行DPCM编码产生的预测误差
- 假设量化误差为0，即 $\dot{e}_n = e_n$
- 定义下列4个预测器，并使用其中1个：

$$\hat{f}(x,y) = 0.97 f(x,y-1)$$

$$\hat{f}(x,y) = 0.5 f(x,y-1) + 0.5 f(x-1,y)$$

$$\hat{f}(x,y) = 0.75 f(x,y-1) + 0.75 f(x-1,y) - 0.5 f(x-1,y-1)$$

$$\hat{f}(x,y) = \begin{cases} 0.97 f(x,y-1) & \text{if } |f(x-1,y) - f(x-1,y-1)| \le |f(x,y-1) - f(x-1,y-1)| \\ 0.97 f(x-1,y) & \text{其它} \end{cases}$$

水平梯度 垂直梯度

例：预测技术的对比



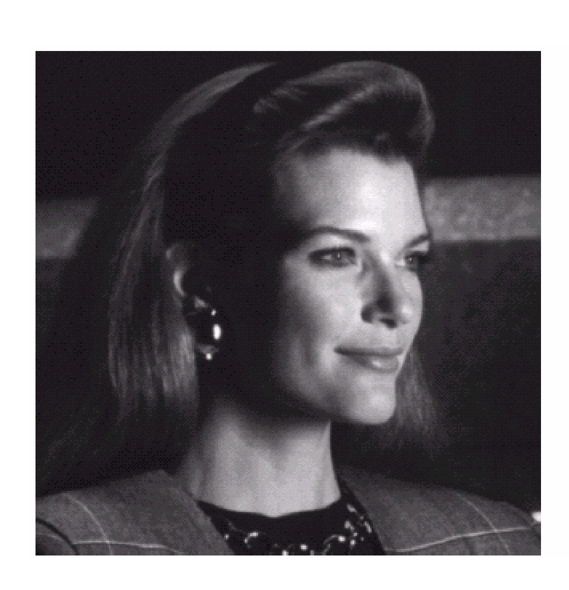


FIGURE 8.23 A 512 × 512 8-bit monochrome image.

例：预测技术的对比



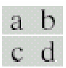
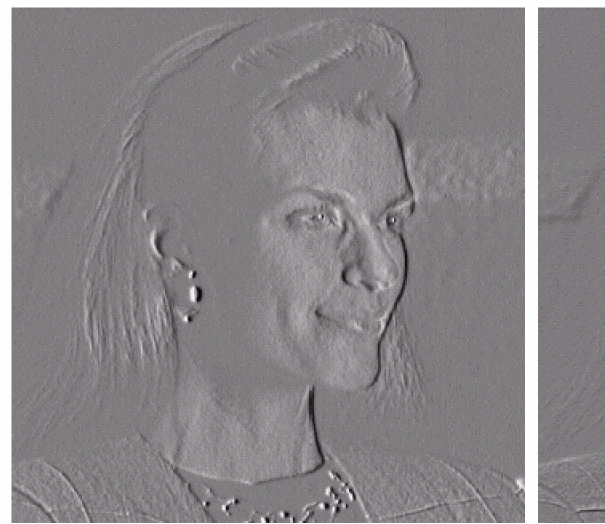


FIGURE 8.24 A comparison of four linear prediction techniques.

1阶

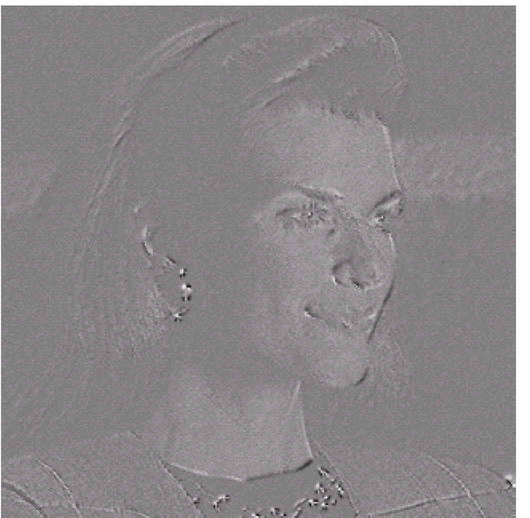




2阶

3阶





4阶

结论：随着预测器阶数的增加误差减少了

有损压缩



- 最优量化

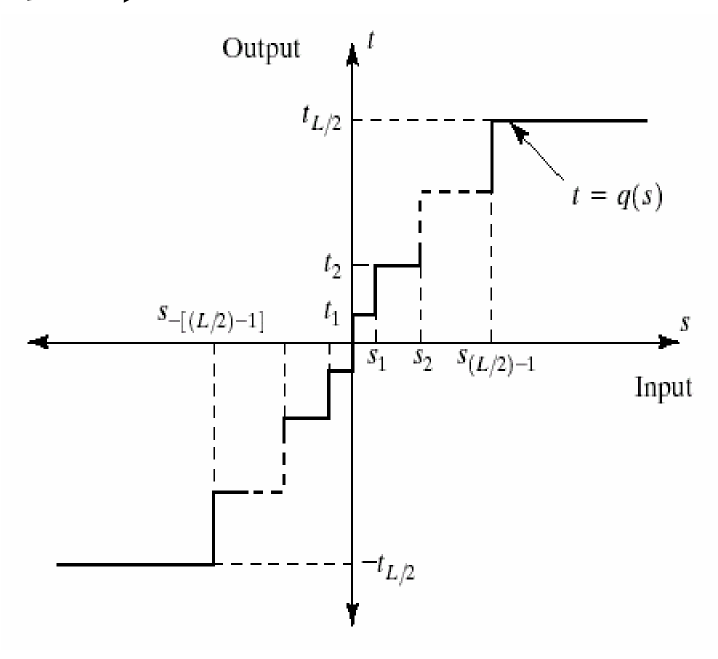


FIGURE 8.25 A typical quantization function.

✓ t=q(s)是s的奇函数

✓ 断点定义了函数的不连续性，被称为量化器的判决和重构级

有损压缩



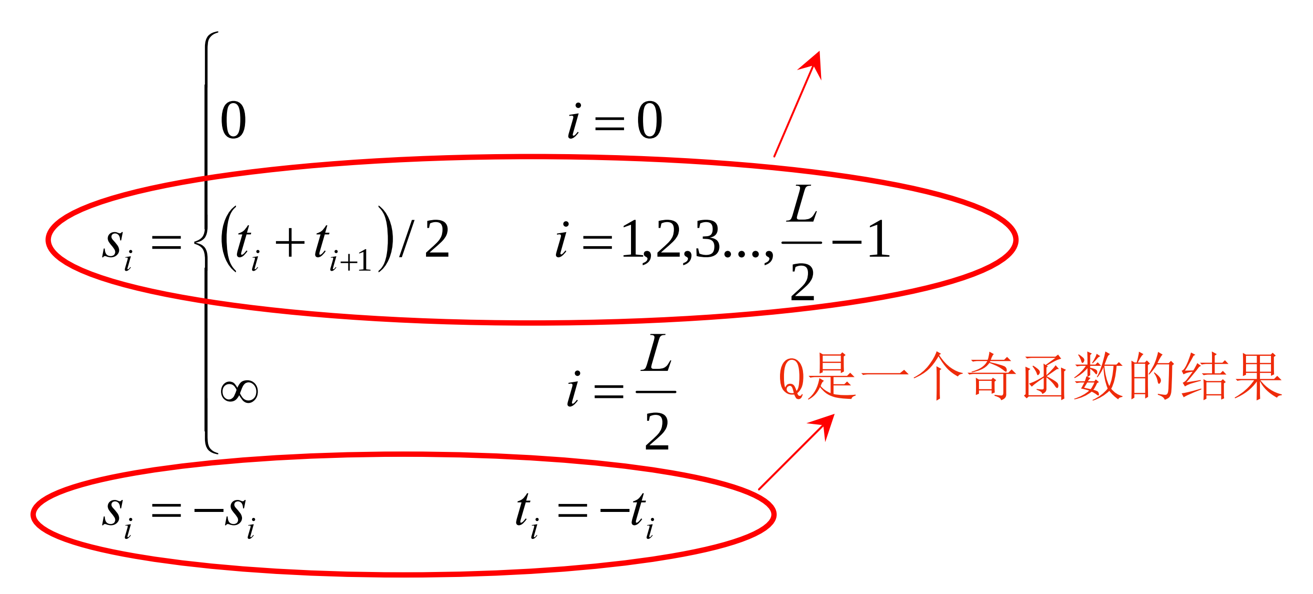
- 最优量化
  - 量化器的设计就是要在给定优化准则和输入概率密度函数p(s)的条件下选择最优的s_i和t_i，优化准则可以是统计的或心理视觉的准则
  - 如果用最小均方量化误差 $E\left[(s-t_i)^2\right]$ 作为准则，且p(s)是个偶函数，那么最小误差条件为

最优量化



$$\int_{s_{i-1}}^{s_i}(s-t_i)p(s)ds=0$$

其中



$$s_i=\begin{cases}0 & i=0\\ (t_i+t_{i+1})/2 & i=1,2,3\ldots,\frac{L}{2}-1\\ \infty & i=\frac{L}{2}\end{cases}$$

$$s_i=-s_i \qquad t_i=-t_i$$

对于任意的L，满足上列公式的s_i和t_i在均方误差意义下是最佳的；相应的量化器称为L层的Lloyd-Max量化器

最优量化



TABLE 8.10 Lloyd-Max quantizers for a Laplacian probability density function of unit variance.

Levels	2		4		8	
i	s_i	t_i	s_i	t_i	s_i	t_i
1	∞	0.707	1.102	0.395	0.504	0.222
2			∞	1.810	1.181	0.785
3					2.285	1.576
4					∞	2.994
θ	1.414		1.087		0.731	

$$\theta = t_i - t_{i-1} = s_i - s_{i-1}$$

最优量化

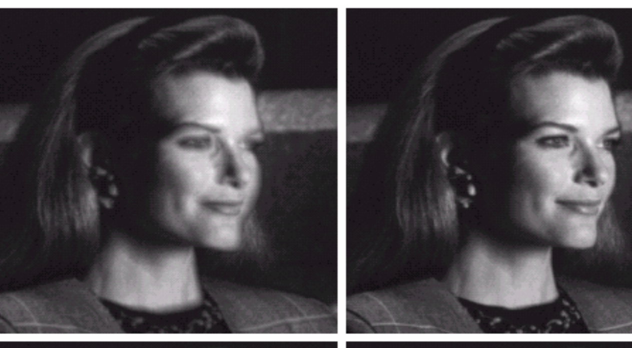


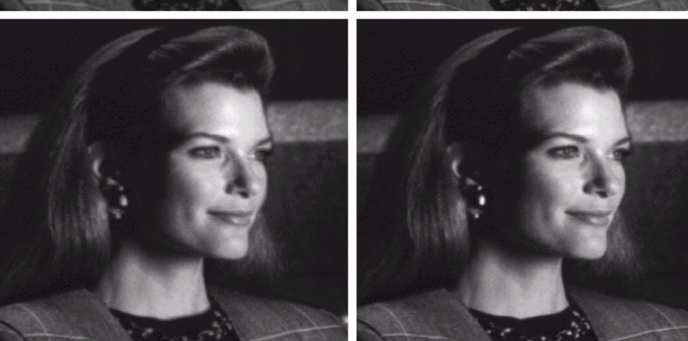
- 总结
  - ✓ 实际应用表明，2级量化器所产生的由于斜率过载而造成的解码图中边缘模糊的程度比4级和8级量化器的程度要高
  - ✓ 具有步长θ的最优均匀量化器在具有相同输出可靠性的条件下能提供比固定长度编码的Lloid-Max量化器更低的码率
  - ✓ Lloid-Max量化器和最优均匀量化器都不是自适应的，但如果根据图像局部性质调解量化值也能提供效率
  - ✓ 可以较细量化缓慢变化区域而较粗量化快速变化区域。这可同时减少颗粒噪声和斜率过载，且码率增加很少
  - ✓ 这也会增加量化器的复杂性

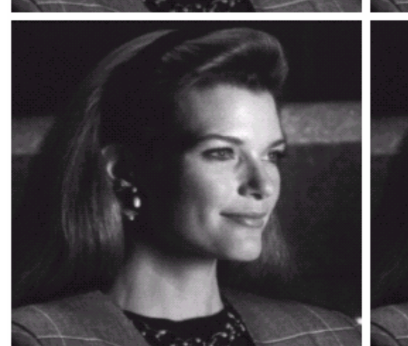
量化和复原例子:DPCM结果图像

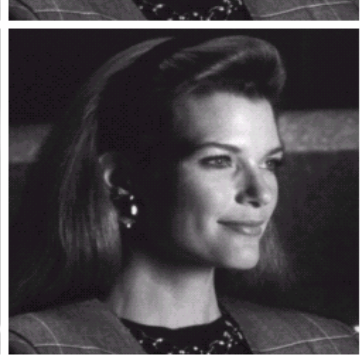


使用3阶预测器 $\hat{f}(x,y)=0.75f(x,y-1)+0.75f(x-1,y)-0.5f(x-1,y-1)$









2层Lloyd-Max量化器

4层Lloyd-Max量化器

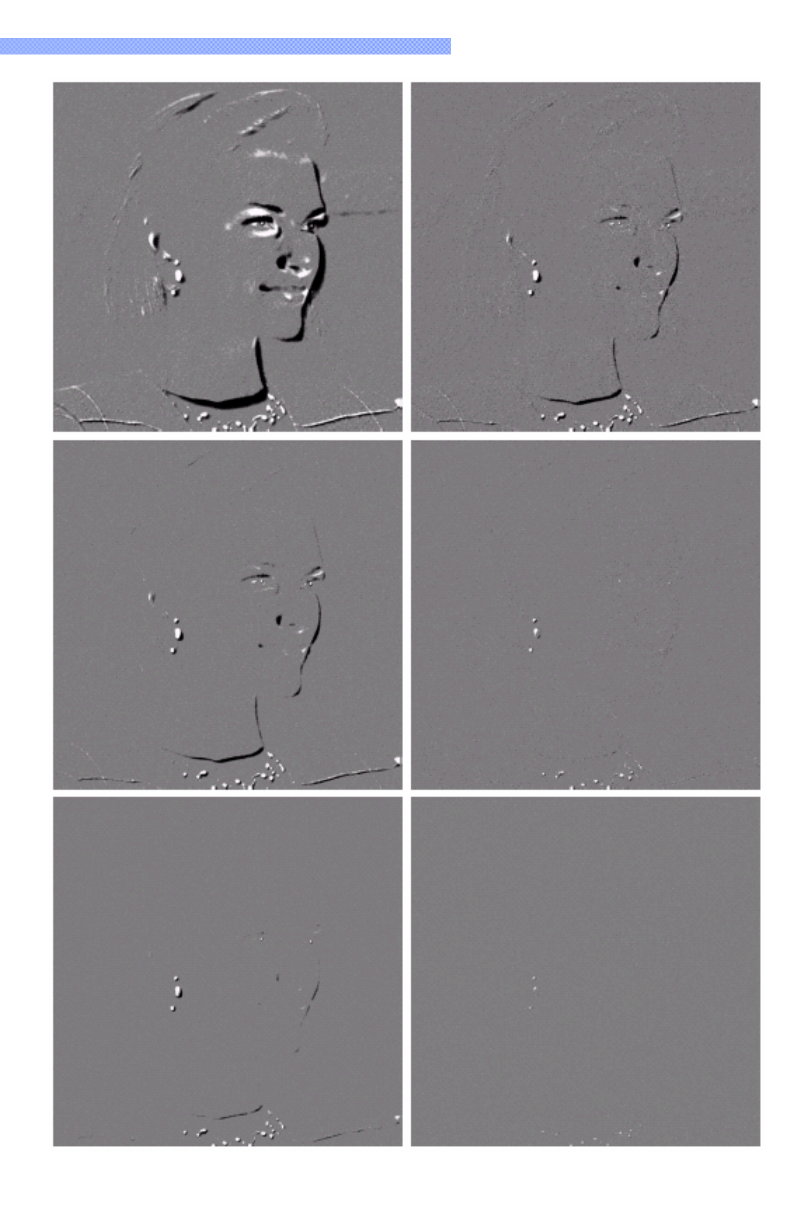
8层Lloyd-Max量化器

自适应量化方法：

在4个可能的量化器

中选择最好的一个

量化和复原例子：标定过的DPCM结果图像



量化和复原例子：标定过的DPCM结果图像

	图a	图b	图c	图d	图e	图f
	Lloyd-Max Quantizer			**Adaptive Quantizer**		
Predictor	**2-level**	**4-level**	**8-level**	**2-level**	**4-level**	**8-level**
Eq. (8.5-16)	30.88	6.86	4.08	7.49	3.22	1.55
Eq. (8.5-17)	14.59	6.94	4.09	7.53	2.49	1.12
Eq. (8.5-18)	9.90	4.30	2.31	4.61	1.70	0.76
Eq. (8.5-19)	38.18	9.25	3.36	11.46	2.56	1.14
Compression	8.00:1	4.00:1	2.70:1	7.11:1	3.77:1	2.56:1

TABLE 8.11 Lossy DPCM root-mean-square error summary.

- ✓ 图a到f中差异图像的rms误差
- ✓ 注明了不同的预测器和量化器
- ✓ 2层自适应量化器的均方误差与4层Lloyd-Max量化器相近
- ✓ 4层自适应量化器的均方误差小于8层Lloyd-Max量化器

有损压缩



- 变换编码
  - ✓ 基于图像变换的编码方法
  - ✓ 用可逆的线性变换（如傅里叶变换）将图像映射成1组变换系数，然后将这些系数量化和编码
  - ✓ 大多数图像变换得到的系数值都很小，这些系数可以较粗地量化，或忽略不计
  - ✓ 虽然失真很小，信息仍然不能完全复原，所以还是有损压缩

变换编码系统



子图分解　　变换　　量化　　编码

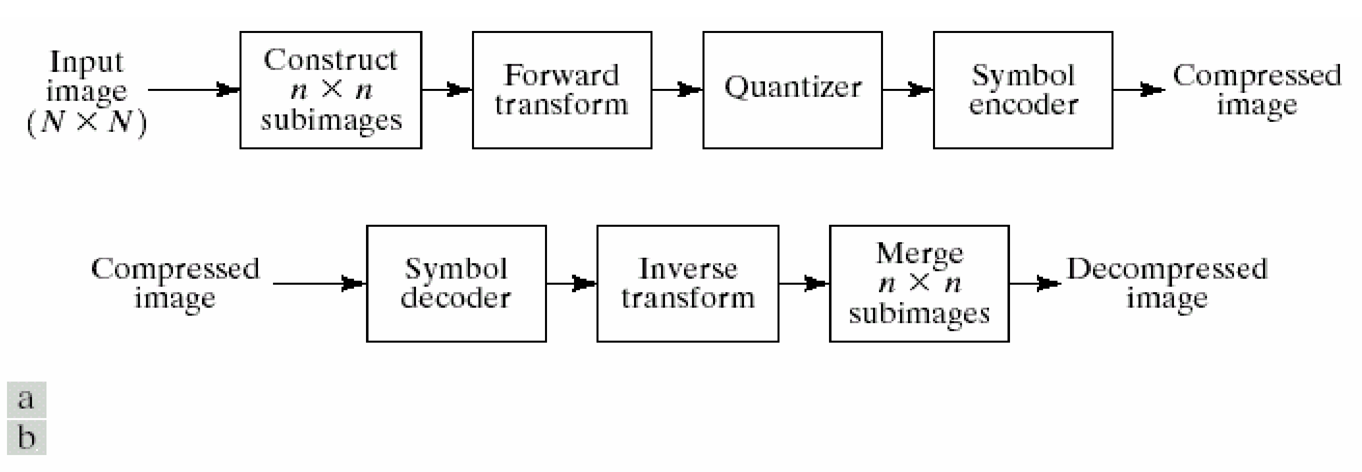


FIGURE 8.28 A transform coding system: (a) encoder; (b) decoder.

变换选择



- 考虑大小为N×N的图像f(x, y)，该图像的正向离散变换T(u, v)表示为

$$T(u,v)=\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)g(x,y,u,v)$$

u, v=0, 1, 2, …, N-1。给定T(u, v), f(x, y)可用离散反变换得到

$$f(x,y)=\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}T(u,v)h(x,y,u,v)$$

x, y=0, 1, 2, …, N-1

g(x, y, u, v)和h(x, y, u, v)分别称为正向和逆向变换核函数

变换选择



- 如果下列等式成立，正向核函数g(x, y, u, v)称为是可分离的

$$g(x,y,u,v)=g_1(x,u)g_2(y,v)$$

如果g_1在函数上等于g_2，则核函数是对称的

$$g(x,y,u,v)=g_1(x,u)g_1(y,v)$$

如果上述等式对逆向核函数h(x, y, u, v)成立，则h(x, y, u, v)也具有上述结论

变换选择



- DFT变换（离散傅里叶变换的简化版本（M=N））

$$g(x,y,u,v)=\frac{1}{N^2}e^{-j2\pi(ux+vy)/N}$$

$$h(x,y,u,v)=e^{j2\pi(ux+vy)/N}$$

- WHT变换

$$g(x,y,u,v)=h(x,y,u,v)=\frac{1}{N}(-1)^{\sum_{i=0}^{m-1}\left[b_i(x)p_i(u)+b_i(y)p_i(v)\right]}$$

WHT变换



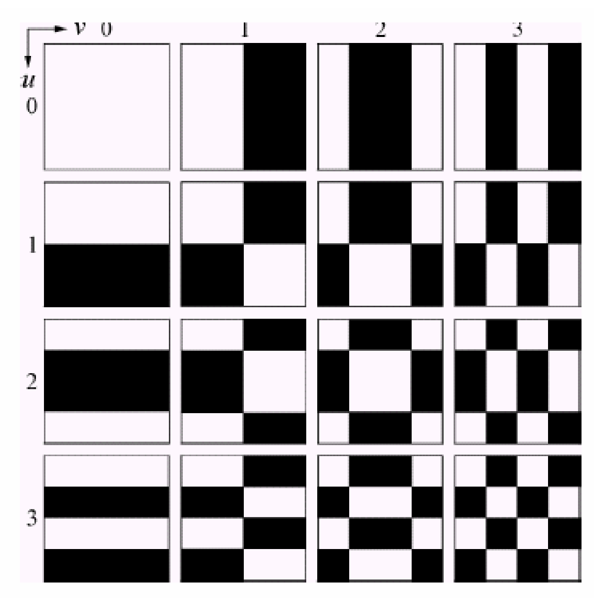


FIGURE 8.29 Walsh-Hadamard basis functions for $N = 4$. The origin of each block is at its top left.

- WHT的重要意义：所有核的值均为+1或-1，实现上简单
- N=4时的核，每个块包括4×4＝16个子方块
- 白色表示＋1，而黑色表示－1
- 左上角的块：u=v=0, x, y=0, 1, 2, 3时g(x, y, 0, 0)的值：＋1
- 左上第二块：u=0, v=1, x, y=0, 1, 2, 3时g(x, y, 0, 1)的值

变换选择



- DCT变换（离散余弦变换）
  - ✓ 一维DCT变换及反变换

$$T(u) = \alpha(u)\sum_{x=0}^{N-1} f(x)\cos\left[\frac{(2x+1)u\pi}{2N}\right] \qquad u = 0,1,\dots, N-1$$

$$f(x) = \sum_{u=0}^{N-1} \alpha(u)T(u)\cos\left[\frac{(2x+1)u\pi}{2N}\right] \qquad x = 0,1,\dots, N-1$$

$$\alpha(u) = \begin{cases} \sqrt{1/N} & \text{当} u = 0 \\ \sqrt{1/N} & \text{当} u = 1,2,\dots, N-1 \end{cases}$$

变换选择



- DCT变换（离散余弦变换）
  - ✓ 二维DCT变换及反变换

$$T(u,v)=\alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cos\left[\frac{(2x+1)v\pi}{2N}\right]$$

$$u,v=0,1,\ldots,N-1$$

$$f(x,y)=\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\alpha(u)\alpha(v)T(u,v)\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cos\left[\frac{(2x+1)v\pi}{2N}\right]$$

$$x,y=0,1,\ldots,N-1$$

DCT变换



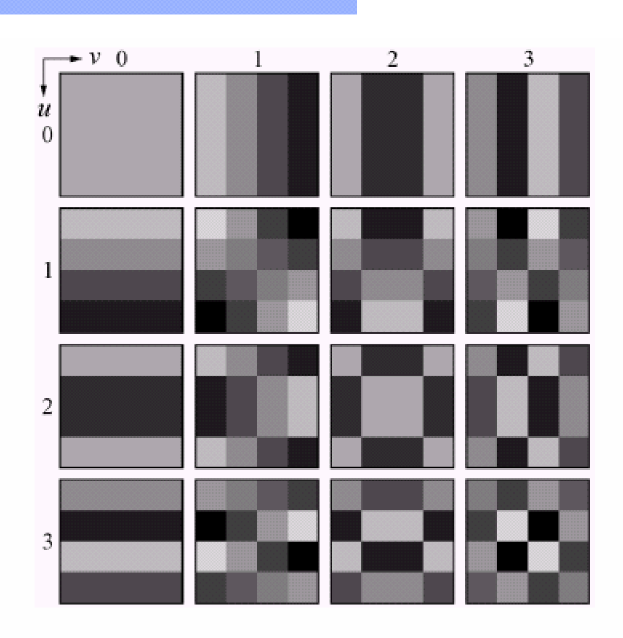


FIGURE 8.30 Discrete-cosine basis functions for $N = 4$. The origin of each block is at its top left.

- g的值不是整数
- 0为黑，1为白，较大的g值，对应较浅的灰度级

三种变换举例



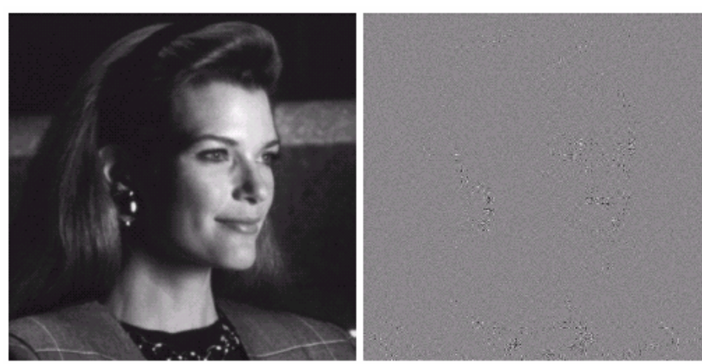
- 说明：
  - ✓ 512×512大小的单色图像
  - ✓ 先将原图分割为8×8大小的子图像，然后用DFT,WHT和DCT的一种，表示每一个子图像
  - ✓ 将得到所有系数的50%去掉，即丢掉32个系数
  - ✓ 对截取的系数阵列进行逆变换，保留32个系数
  - ✓ 丢掉的32个系数，对复原图像质量的视觉影响很小
  - ✓ 产生的均方误差

$$rms_{DFT} > rms_{WHT} > rms_{DCT}$$

三种变换举例

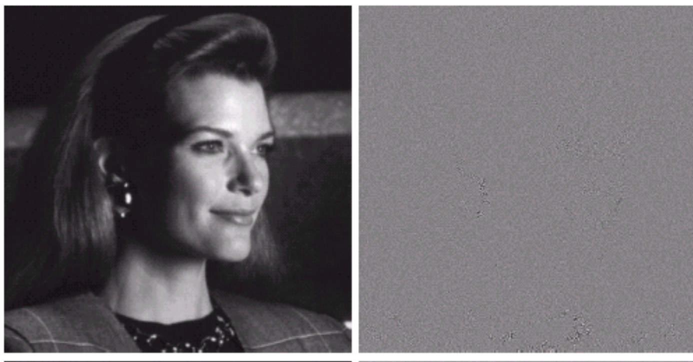


DFT还原图像



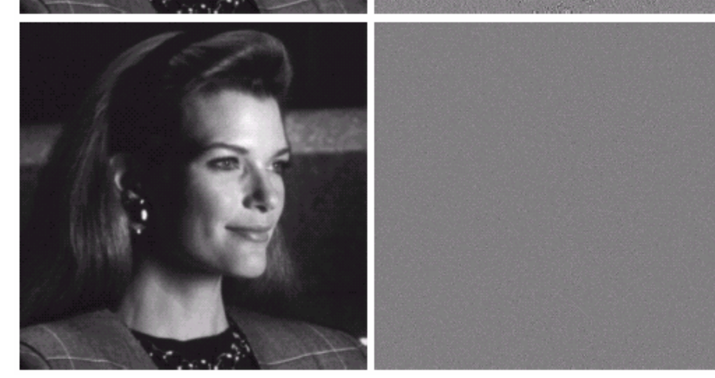
rms_{DFT}=1.28

WHT还原图像



rms_{WHT}=0.86

DCT还原图像



rms_{DCT}=0.68

三种变换举例

- 结论
  - ✓ DCT的信息压缩能力比DFT和WHT的能力要强
  - ✓ WHT是最容易实现的
  - ✓ DCT在信息压缩能力和计算复杂性之间提供了很好的平衡，因此，许多变换编码系统都是以DCT变换为基础的
  - ✓ 对比其它方法，DCT变换具有使用单一的集成电路就可以实现，可以将最多的信息包装在最少的系数之中
  - ✓ 可使“分块噪声”的块效应最小，这些分块噪声是由于图像之间的可见边界造成的

三种变换举例



- 结论（续）
  - ✓ DFT的n点周期性造成边界中断，当DFT系数被截尾或量化时，边界值以分块噪声的形式出现，即相邻子图像之间的边界变得可见。而DCT变换不会产生固有的边界中断

DFT

DCT

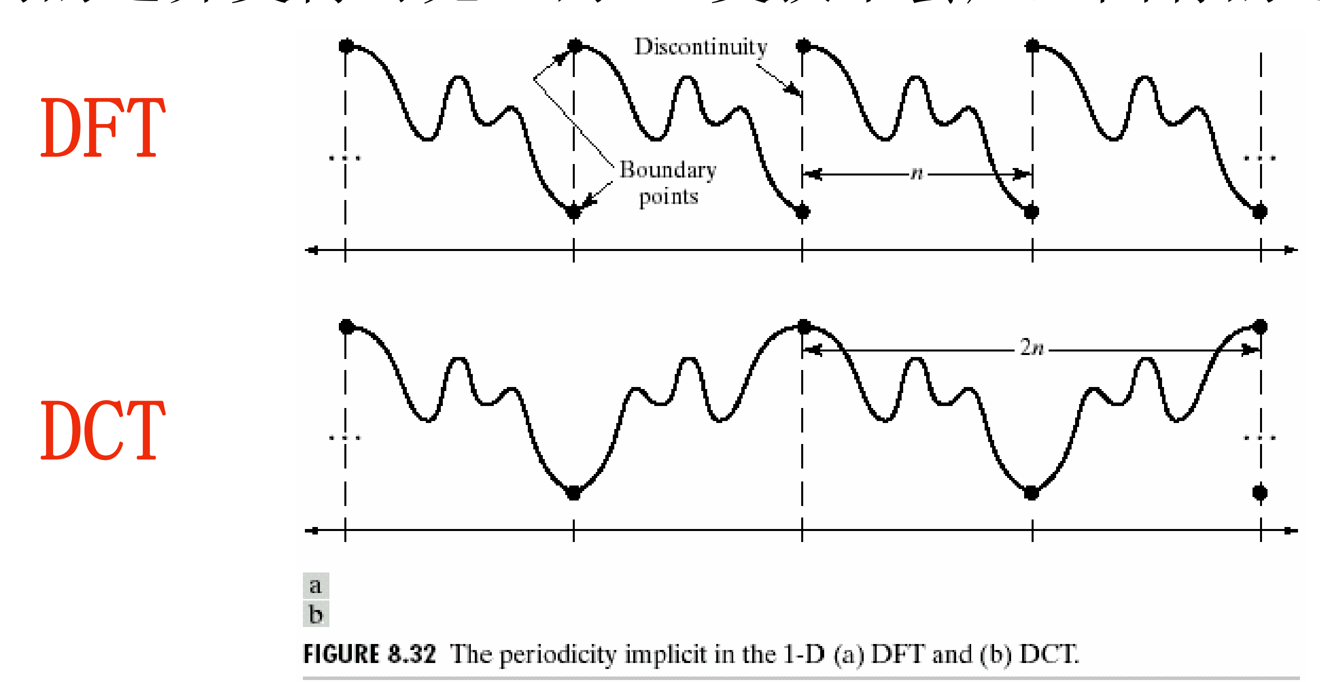


FIGURE 8.32 The periodicity implicit in the 1-D (a) DFT and (b) DCT.

变换编码



- 子图像尺寸选择
  - ✓ 子图像尺寸是影响变换编码误差和计算复杂度的一个重要因素
  - ✓ 一般满足下列2个条件
    - ➢ 相邻子图像之间的相关（冗余）减少到可接受
    - ➢ 子图像的长和宽都是2的整数次幂。这主要是为了简化对子图像变换的计算
  - ✓ 最常用的子图像尺寸是8×8和16×16

子图像尺寸选择



- 一个试验例子：

✓ 把图像分割为n×n的子图像(n=2,4,8,16和32)

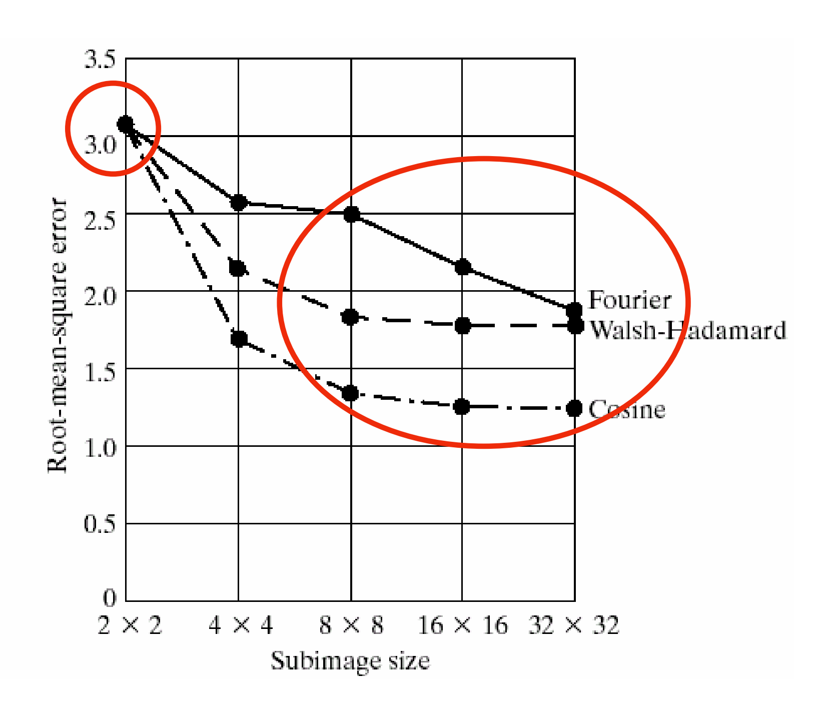
✓ 计算每幅子图像的变换，截取75%的系数，对截取后的阵列进行反变换得到

✓ 当图像尺寸大于8×8，WHT和DCT曲线变得比较平缓，而DFT曲线下降得较快。对更大的n值，DFT将穿过WHT并逼近DCT

✓ 当图像尺寸为2×2时，3条曲线交于同1点。此时各变换都只保留了4个系数中的1个（25%），即直流分量

子图像尺寸选择


PEKING UNIVERSITY


FIGURE 8.33
Reconstruction error versus subimage size.
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Root-mean-square error
2 × 2
4 × 4
8 × 8
16 × 16
32 × 32
Subimage size
Fourier
Walsh-Hadamard
Cosine

变换编码



- 比特分配
  - ✓ 截断误差和2个因素有关
    - ➢ 截除的变换系数的数量和相对重要性
    - ➢ 用来表示所保留系数的精度
  - ✓ 在多数变换编码系统中，保留的系数是根据下列2个准则之一来选择确定的
    - ➢ 最大方差准则，称为分区编码
    - ➢ 最大幅度准则，称为阈值编码
  - ✓ 整个对变换子图像的系数截断、量化和编码的全过程称为比特分配

比特分配效果举例



- 对每个8×8子图像丢弃87.5%的DCT系数的两种近似

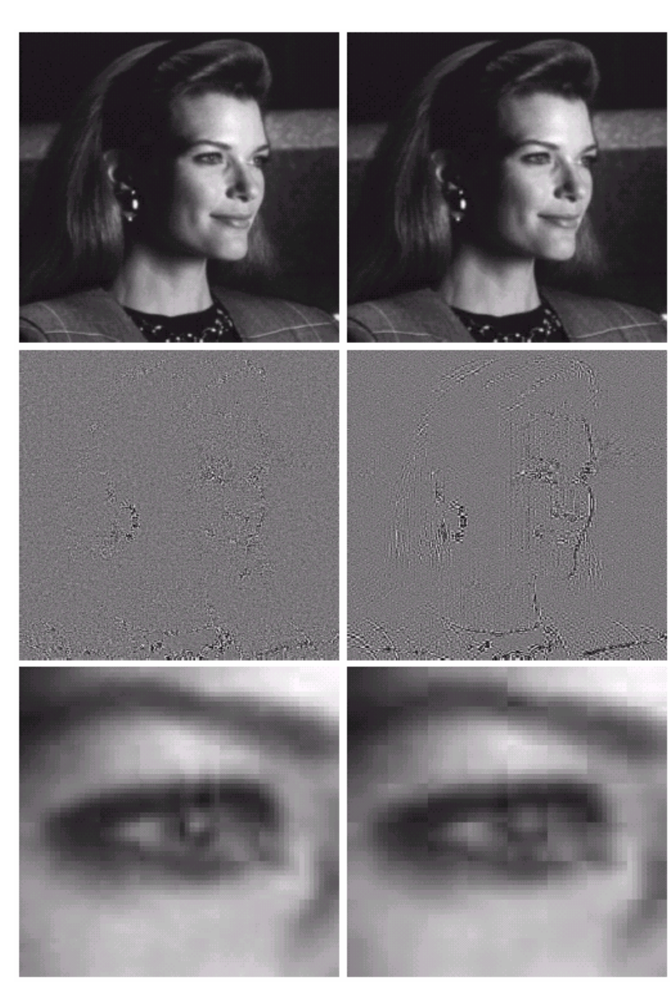
阈值编码：

保留8个最大的

变换系数，误差

比分区编码小

(64-8)/64=87.5%



分区编码

变换编码



- 其它编码方法
  - ✓ 子带(subband)编码
  - ✓ 小波编码
  - ✓ 分形（fractal）编码
  - ✓ 矢量量化(vector quantization,VQ)编码
  - ✓ 统计编码/熵编码

图像压缩标准



- 相关的国际组织
  - ✓ ISO（International Standardization Organization,国际标准化组织）
  - ✓ ITU（International Telecommunication Union,国际电信联盟），前身是CCITT（国际电话电报咨询委员会）
- 相关工作
  - ✓ 覆盖了从二值到灰度（彩色）值的静止和运动图像
  - ✓ 采用的大部分基本技术前面已经介绍，主要包括预测和变换编码技术

图像压缩标准



- 可分成下面几个系列
  - ✓ 用于压缩二值图像
  - ✓ 用于压缩静止图像，包括灰度和彩色图像
  - ✓ 用于压缩运动图像（视频），包括灰度和彩色图像
- 上述2个组织还在制定一些新的标准，其中一些已经超出纯图像编码的范围

二值图像压缩标准



- G3和G4
  - ✓ 这2个标准是由CCITT的两个小组(Group 3 和 Group 4)负责制定的
  - ✓ 最初是CCITT为传真应用而设计的，现也用于其它方面
  - ✓ G3采用了非自适应、1维行程编码技术。对每组N行（N=2或N=4）扫描线中的后N-1行也可以用2维方式编码
  - ✓ G4是G3的1种简化版本，其中只使用2维编码

二值图像压缩标准



- G3和G4（续）
  - ✓ CCITT在制定标准期间曾选择了1组共8幅具有一定代表性的“试验”图用来评判各种压缩方法
  - ✓ 它们既包括打印的文字，也包括用几种语言手写的文字，另外还有少量的线绘图
  - ✓ G3对它们的压缩率约为15:1
  - ✓ G4的压缩率一般比G3高1倍

二值图像压缩标准



- JPEG(Joint Picture Expert Group)
  - ✓ 由ISO和CCITT联合成立的专家组负责制定静态图像（彩色与灰度图像）的压缩算法
  - ✓ 定义了3种编码系统
    - ➢ 基于DCT的有损编码基本系统，可用于绝大多数压缩应用场合
    - ➢ 用于高压缩比、高精确度或渐进重建应用的扩展编码系统
    - ➢ 用于无失真应用场合的无损系统

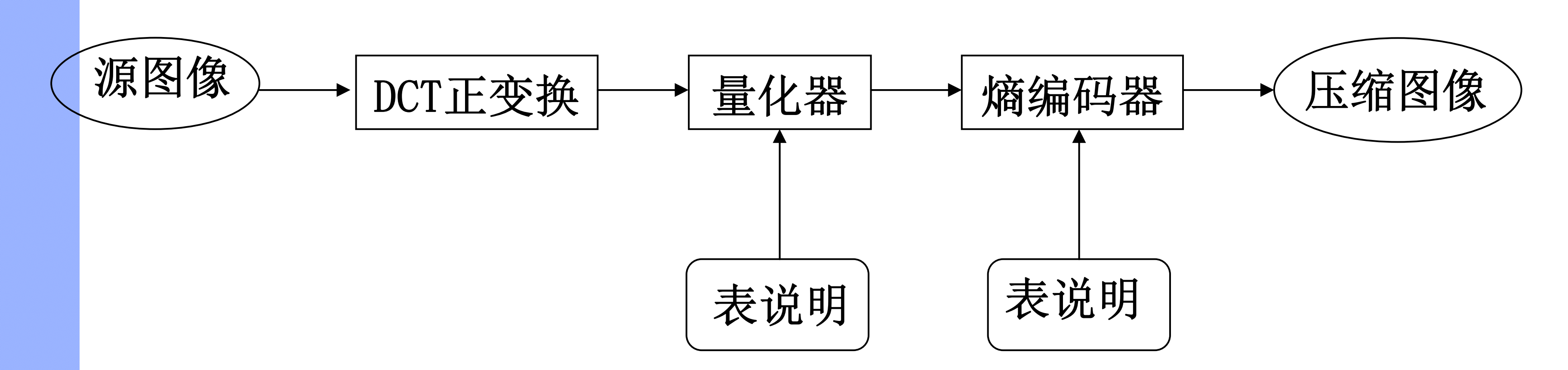
二值图像压缩标准



- JPEG(Joint Picture Expert Group)
  - ✓ JPEG对录像机质量的静止图像的压缩率一般可达到25:1

JPEG基本系统编码器框图





- 压缩过程
  - ✓ DCT计算
  - ✓ 量化
  - ✓ 变长码赋值

JPEG基本系统编码器框图



- 具体压缩过程
  - ✓ 先把整个图像分解成多个8×8的图像块；
  - ✓ 8×8的图像块经过DCT变换后，低频分量都集中在左上角，高频分量则分布在右下角（DCT变换类似于低通滤波器），因为低频分量包含了图像的主要信息，所以可以忽略高频分量，达到压缩的目的；
  - ✓ 使用量化操作去掉高频分量，量化操作就是将某一个值除以量化表中的对应值。由于量化表中左上角的值较小，而右下角的值较大，这样达到保持低频分量，抑制高频分量的目的；
  - ✓ 在左上角的低频分量中，F(0,0)代表了直流(DC)系数，即8×8子块的平均值。由于两个相邻图像块的DC系数相差很小，所以采用差分编码DPCM，其它63个元素是交流（AC）系数，采用之字型（zig-zag）顺序进行行程编码，使系数为0的值更集中；

JPEG基本系统编码器框图



- 具体压缩过程

  ✓ 在得到DC码字和AC行程码字后，为了进一步提高压缩比，再进行熵编码，采用了Huffman编码

视频压缩标准



- 视频压缩标准
  - ✓ 基本概念
  - ✓ 电视会议标准
    - ➢ H.261、H.263
  - ✓ MPEG标准
    - ➢ MPEG-1、MPEG-2、MPEG-4、MPEG-7和MPEG-21

一个运动图像例子



DEMO

观察一段视频解压缩后的图像

MPEG帧的分类



- I帧（Intra-picture）
  不需要参考其它画面而独立进行压缩编码的画面
- P帧（Predicted-picture）
  参考前面已编码的I或P画面进行预测编码的画面
- B帧（Bidirectional-picture）
  既参考前面的I或P画面、又参考后面的I或P画面进行双向预测编码的画面

视频压缩标准

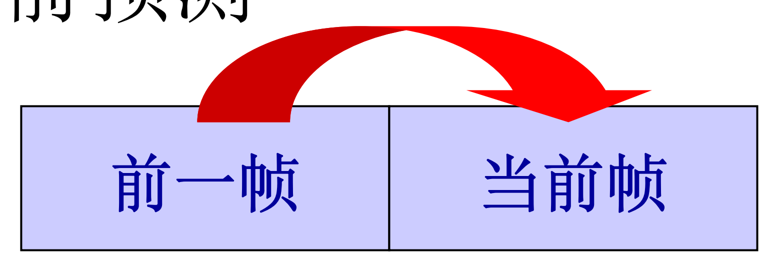


- 连续帧图像压缩的基本思想
  - ✓可以根据同帧附近像素来加以预测，被称为：帧内编码技术
  - ✓可以根据附近帧中的像素来加以预测，被称为：帧间编码技术

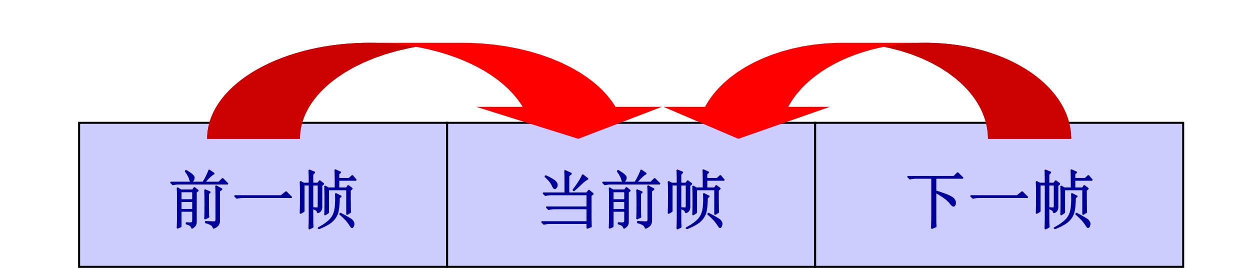
视频压缩标准



- 帧间预测编码
  - ✓ 向前预测



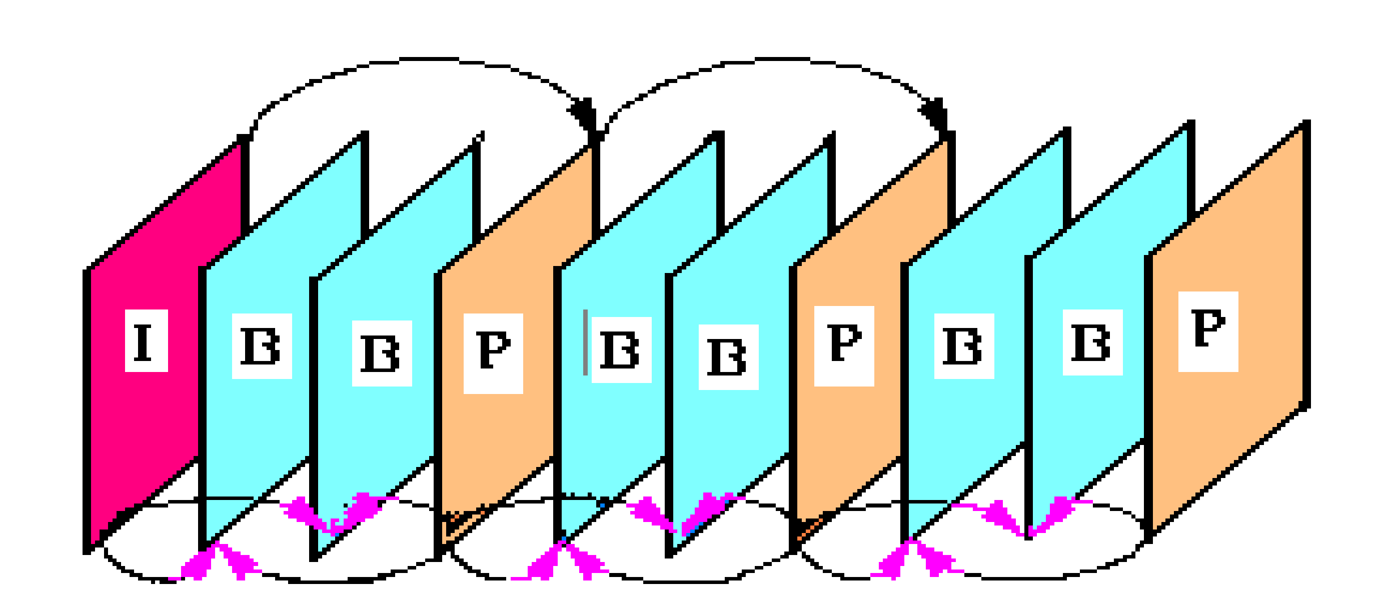
  - ✓ 双向预测



MPEG帧的分类



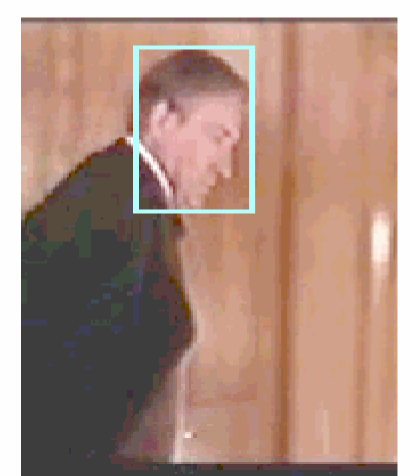
MPEG 的画面组(GOP)



MPEG帧的分类

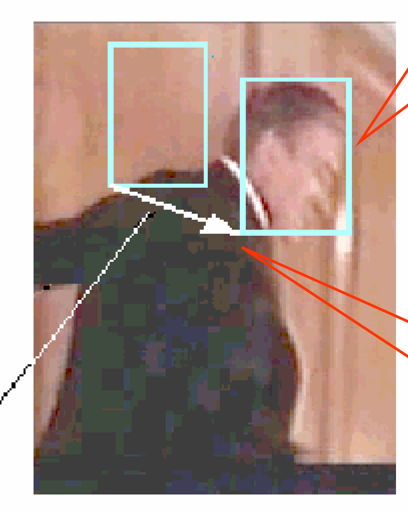


《泰坦尼克》电影上的镜头



时刻1

运动矢量



时刻2

由于画面内容有连贯性,因此当前画面的内容可以看作是前面画面内容的位移(运动)。

画面各处相应的位移量(幅度、方向)各有不同，因此画面应细分成块(宏块)，以宏块为单位，找出两帧画面中相应宏块之间的位移量---运动矢量

画面的重新排序



例

GOP(Group of Pictures)

0	1	2	3	4	5	6	7	8	9	10
B	B	I	B	B	P	B	B	P	B	I

- 画面的显示顺序是：　B B I B B P B B P B I
  0 1 2 3 4 5 6 7 8 9 10
- 画面的编码顺序是：　I B B P B B P B B I B
  2 0 1 5 3 4 8 6 7 10 9

因此，每一个GOP中的画面在编码前和解码后都必须重排序

MPEG帧的分类



DEMO

观察MPEG的运动矢量

ISO国际标准的制定过程



- 工作组草案(WD)=>委员会草案(CD)=>国际标准草案(DIS)=>国际标准(IS)
  - ✓ 工作组草案WD(Working Draft)：工作组WG(Working Group)准备的工作文件
  - ✓ 委员会草案CD(Committee Draft)：从WD提升上来的文件。这是ISO文档的最初形式，它由ISO内部正式调查研究和投票表决。
  - ✓ 国际标准草案DIS(Draft International Standard)：成员国对CD的内容和说明满意之后由CD提升上来的文件。
  - ✓ 国际标准IS(International Standard)：由成员国、ISO的其他部门和其他委员会投票通过之后出版发布的文件。

电视会议标准



- H.261标准
  - ✓ 应用范围：电视会议
  - ✓ 主要编码技术：
    - ➢ DCT变换
    - ➢ 向前运动补偿预测
    - ➢ Zig-zag排序
    - ➢ 霍夫曼编码

IPPPPPPIPPP.....

电视会议标准



- H.263标准
  - ✓ 应用范围：可视电话
  - ✓ 主要编码技术：
    - ➢ DCT变换
    - ➢ 双向运动补偿预测
    - ➢ Zig-zag排序
    - ➢ 霍夫曼编码

IBBPBBPBBIBBP . . . .

多媒体压缩标准



- MPEG系列标准(Moving Picture Expert Group)

  - ✓ ISO/IEC/JTC1/SC29的一个工作组WG11，1988年成立，目前有25个国家(团体)的200多个公司300多名成员分10个组工作
  - ✓ JPEG的目标是专门集中于静止图像压缩，MPEG的目标是针对运动图像的数据压缩，但是JPEG和MPEG有密切联系
  - ✓ MPEG专家小组，不仅限制于数字视频压缩，音频及音频和视频的同步问题都不能脱离视频压缩独立进行

MPEG-1标准



- MPEG-1标准(ISO/IEC11172). 1992年发布。

  用于1.5Mbps数据传输率的运动图像及其伴音的编码。主要应用于 VCD，MP3音乐等

- MPEG-1在JPEG和H.261等优秀标准的基础上，对参加竞争的14个方案,通过反复协调而得到统一,从而成为先进、合理、质量高、成本低的优秀标准

- MPEG-1促进了大规模集成电路专用芯片的发展,为多媒体技术和产品的繁荣立下了功劳

MPEG-1标准（续）



- MPEG-1标准
  - ✓ 主要编码技术：
    - DCT变换
    - 前向、双向运动补偿预测
    - Zig-zag排序
    - 霍夫曼编码、算术编码
    - 每15帧至少要有一个I帧

IBBPBBPBBIBBP．．．．

MPEG-2标准



- MPEG-2标准(ISO/IEC13818). 1994年发布
- 主要针对高清晰度电视(HDTV) 的视频及伴音信号，典型传输速率为10Mbps，与MPEG-1兼容，适用于1.5Mbps～60Mbps甚至更高速率的编码范围

MPEG-2标准（续）



- 以MPEG-2作为视音频压缩标准的数字卫星电视接收机已在欧美形成了很大市场，美国高级电视联盟(ATV Grand Alliance)和欧洲数字视频广播计划(Digital Video Broadcast Project)先后决定将MPEG -2用于高清晰度电视(HDTV)广播中
- 新一代的数字视盘DVD采用MPEG-2作为其视音频压缩标准(注:欧、美 、日在视频方面采用MPEG-2标准,而在音频方面则采用AC-3标准)

MPEG-4标准



- MPEG-4 标准(ISO/IEC 14496) .
  - ✓ 1999年5月形成国际标准(版本1)，是一种基于对象的视(音)频编码标准,目标是支持各种多媒体应用(主要侧重于对多媒体信息内容的访问)
  - ✓ 为了适应在窄带宽(一般指<64kbit/s)通信线路上对动态图像进行传输的要求
  - ✓ 总目标是对各种音频视频，主要包括：静止图像、序列图像、计算机图形、3D模型、动画、语言、声音等进行统一有效的编码

MPEG-4标准（续）



- MPEG-4 标准(ISO/IEC 14496).
  - ✓ 既支持固定码流也支持变码流，对3种码流范围的视频已达最优：
    - < 64k bit/s
    - 64 - 384k bit/s
    - 364 kbit/s – 4M bit/s
    - 目前还在考虑直到50M bit/s的码流

MPEG-7标准



- MPEG-7标准
  - “多媒体内容描述接口”(Multimedia Content Description Interface)的标准
  - 该标准将确定各种类型的多媒体信息的标准描述方法，可应用于数字图书馆、各种多媒体目录服务、广播媒体的选择，以及多媒体编辑等领域
  - 目标是指定1组描述不同多媒体信息的标准描述符，这些描述要与信息内容相关以便能快速和有效地查询各种多媒体信息

MPEG-7标准（续）



- MPEG-7标准
  - ✓ MPEG-7采取的描述方案和方法与被描述内容是否编码或如何存储无关，例如视觉信号仍可以用已有的各种编码方案（如JPEG，MPEG-1,MPEG-2,MPEG-4等）进行编码
  - ✓ 计划2001年9月形成，但还有许多研究工作要做

MPEG-21标准



- MPEG-21标准. (Multimedia Framework)

  多媒体框架标准适合于多种应用领域中各种类型用户对各种类型内容的传送和使用，可通过大范围的网络和设备允许对多媒体资源的透明和增强使用

数字图像处理（9）



任何问题？

研究生课程



数字图像处理
Digital Image Processing

彭宇新

北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

形态学图像处理



- 概述
- 集合论基础知识
- 膨胀和腐蚀：产生滤波器作用
- 开操作和闭操作：产生滤波器作用
- 击中或击不中变换
- 形态学的主要应用：

边界提取、区域填充、连通分量的提取、凸壳、细化、粗化等

形态学图像处理



- 概述
  - ✓ 形态学一般指生物学中研究动物和植物结构的一个分支
  - ✓ 用数学形态学（也称图像代数）表示以形态为基础对图像进行分析的数学工具
  - ✓ 基本思想是用具有一定形态的结构元素去度量和提取图像中的对应形状以达到对图像分析和识别的目的
  - ✓ 形态学图像处理的数学基础和所用语言是集合论

形态学图像处理



- 概述（续）
  - ✓ 形态学图像处理的应用可以简化图像数据，保持它们基本的形状特性，并除去不相干的结构
  - ✓ 形态学图像处理的基本运算有4个：膨胀、腐蚀、开操作和闭操作

形态学图像处理



- 集合论基础知识
  - 复习集合的并、交、补、差

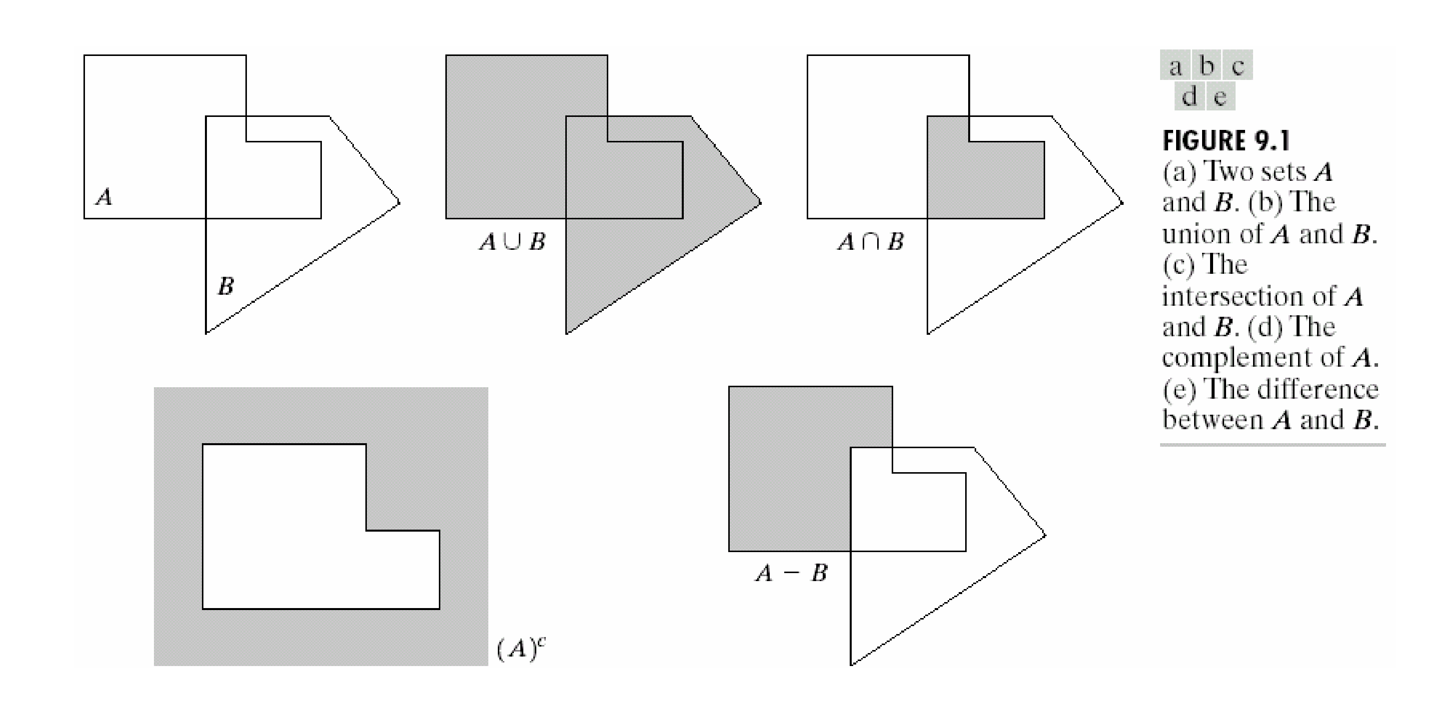


FIGURE 9.1 (a) Two sets A and B. (b) The union of A and B. (c) The intersection of A and B. (d) The complement of A. (e) The difference between A and B.

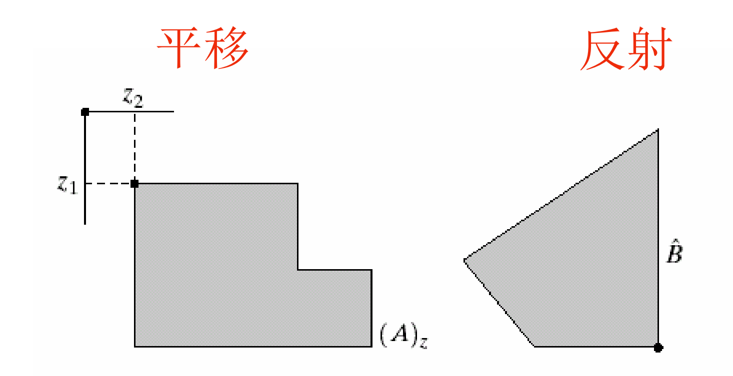
形态学图像处理



- 集合论基础知识
  - ✓ 集合B的反射 $\hat{B}$ ，定义为 $\hat{B}=\{w \mid w=-b, b\in B\}$

    即关于原集合原点对称
  - ✓ 集合A平移到点z=(z_1, z_2)，表示为$(A)_z$，定义为

$$(A)_z=\{c \mid c=a+z, a\in A\}$$



a b

FIGURE 9.2
(a) Translation of *A* by *z*.
(b) Reflection of *B*. The sets *A* and *B* are from Fig. 9.1.

形态学图像处理



- 膨胀和腐蚀
  - ✓ 膨胀：使图像扩大

    A和B是两个集合，A被B膨胀定义为：

$$A \oplus B = \left\{ z \mid \left(\hat{B} \right)_z \cap A \neq \phi \right\}$$

  - 上式表示：B的反射进行平移与A的交集不为空
  - B的反射：相对于自身原点的映象
  - B的平移：对B的反射进行位移

形态学图像处理



- 膨胀和腐蚀
  - ✓ 膨胀的另一个定义

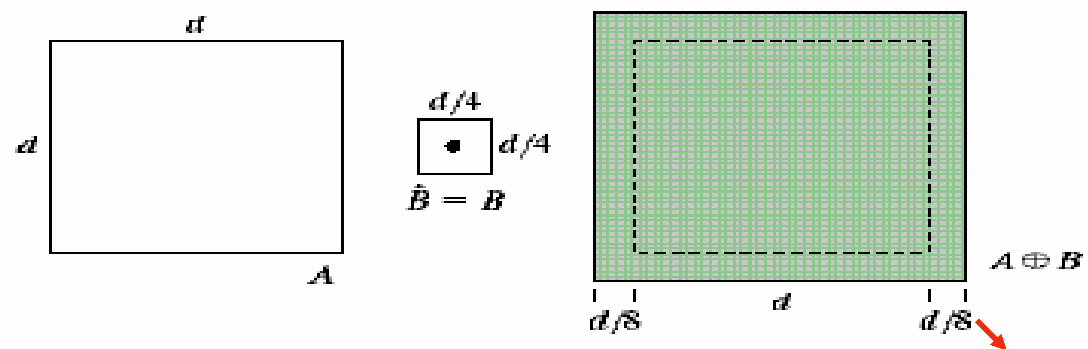
$$A \oplus B = \left\{ z \,\middle|\, \left(\hat{B} \right)_z \cap A \subseteq A \right\}$$

➢ 上式表示：B的反射进行平移与A的交集是A的子集

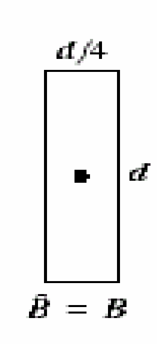
膨胀举例

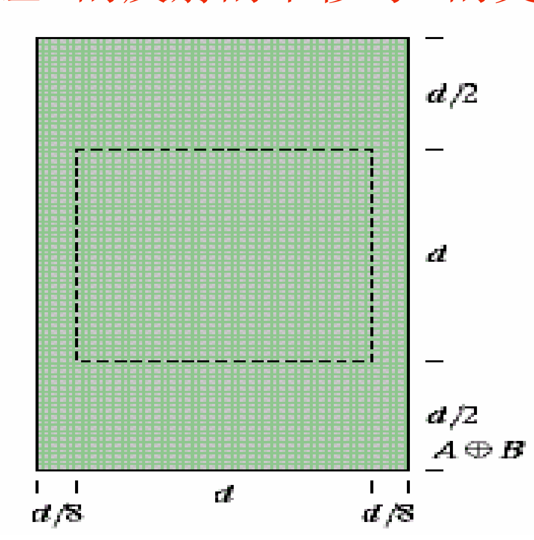


集合B关于原点对称



保证B的反射的平移与A的交集不为空





膨胀应用举例



✓ 桥接文字裂缝

优点：在一幅二值图像中直接得到结果，对比低通滤波方法

Historically, certain computer programs were written using only two digits rather than four to define the applicable year. Accordingly, the company's software may recognize a date using "00" as 1900 rather than the year 2000.

Historically, certain computer programs were written using only two digits rather than four to define the applicable year. Accordingly, the company's software may recognize a date using "00" as 1900 rather than the year 2000.

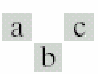


FIGURE 9.5 (a) Sample text of poor resolution with broken characters (magnified view). (b) Structuring element. (c) Dilation of (a) by (b). Broken segments were joined.

0	1	0
1	1	1
0	1	0

→ 膨胀的结构元素B

形态学图像处理



- 膨胀和腐蚀
  - 腐蚀：使图像缩小

    A和B是两个集合，A被B腐蚀定义为：

$$A \Theta B = \{ z \mid (B)_z \subseteq A \}$$

腐蚀举例



A被B腐蚀的结果

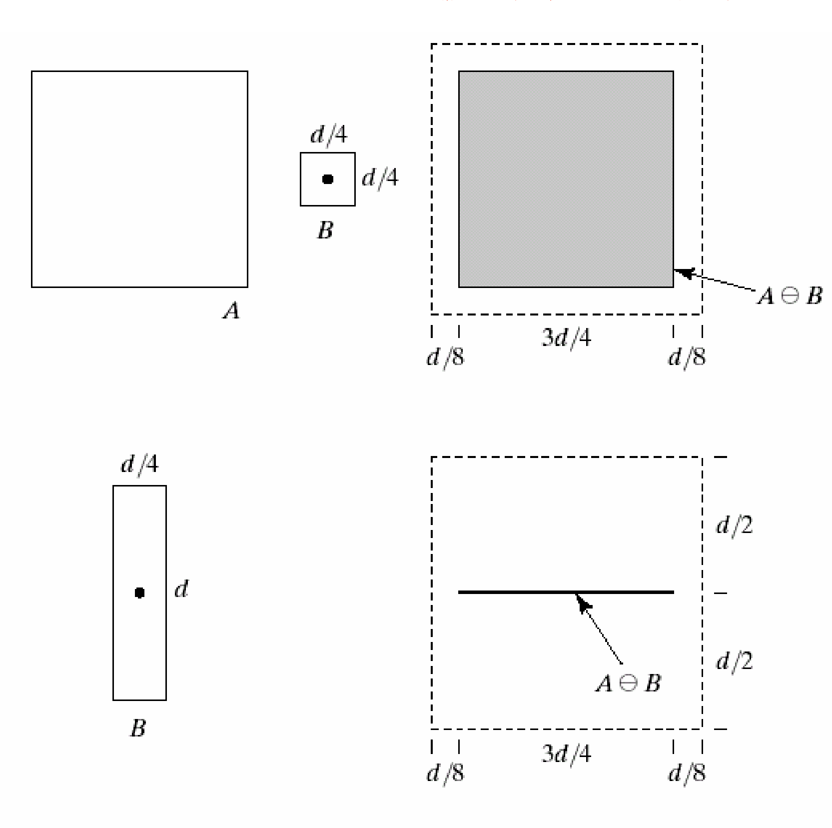


FIGURE 9.6 (a) Set A. (b) Square structuring element. (c) Erosion of A by B, shown shaded. (d) Elongated structuring element. (e) Erosion of A using this element.

腐蚀应用举例



✓ 使用腐蚀消除图像的细节部分，产生滤波器的作用

包含边长为1,3,5,7,9和15像素正方形的二值图像

使用13×13像素大小的结构元素腐蚀原图像的结果

使用13×13像素大小的结构元素膨胀图b，恢复原来15×15尺寸的正方形

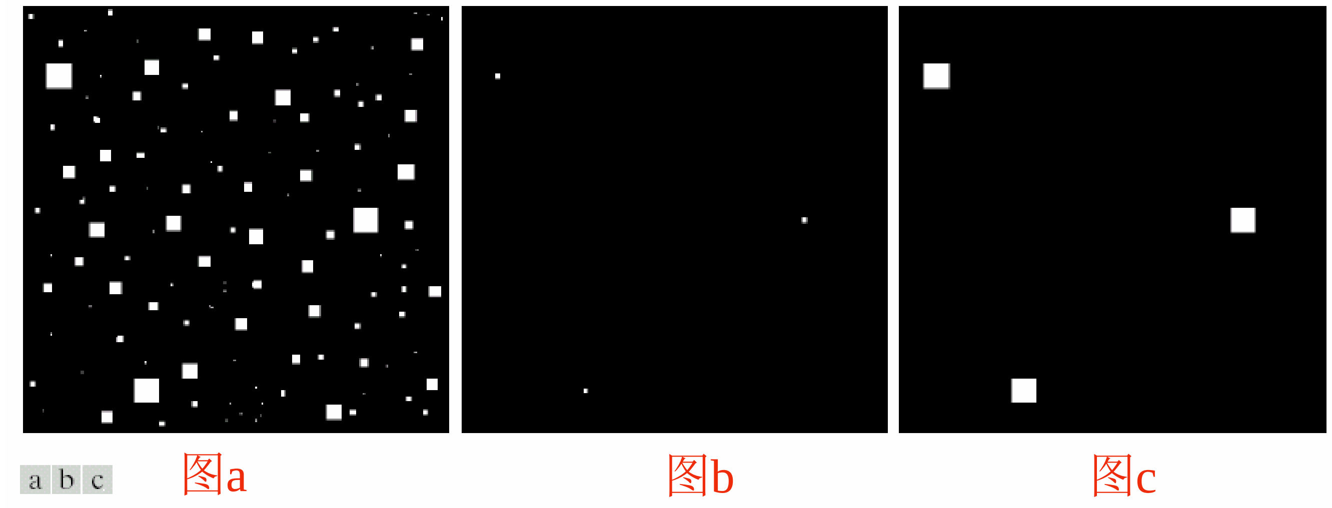


FIGURE 9.7 (a) Image of squares of size 1, 3, 5, 7, 9, and 15 pixels on the side. (b) Erosion of (a) with a square structuring element of 1's, 13 pixels on the side. (c) Dilation of (b) with the same structuring element.

形态学图像处理



- 开操作和闭操作

  ✓ 开操作：使图像的轮廓变得光滑，断开狭窄的间断和消除细的突出物

  ➢ 使用结构元素B对集合A进行开操作，定义为：

  $$A \circ B = (A \Theta B) \oplus B$$

  ➢ 含义：先用B对A腐蚀，然后用B对结果膨胀

  ➢ 另一个定义 $A \circ B = \cup \{ (B)_z \mid (B)_z \subseteq A \}$

形态学图像处理



- 开操作和闭操作

  ✓ 闭操作：同样使图像的轮廓变得光滑，但与开操作相反，它能消除狭窄的间断和长细的鸿沟，消除小的孔洞，并填补轮廓线中的裂痕

  - 使用结构元素B对集合A进行闭操作，定义为：

$$A \bullet B = (A \oplus B) \Theta B$$

  - 含义：先用B对A膨胀，然后用B对结果腐蚀

开操作的几何解释



- ✓ $A \circ B$ 的边界通过B中的点完成
- ✓ B在A的边界内转动时，B中的点所能到达的A的边界的的最远点

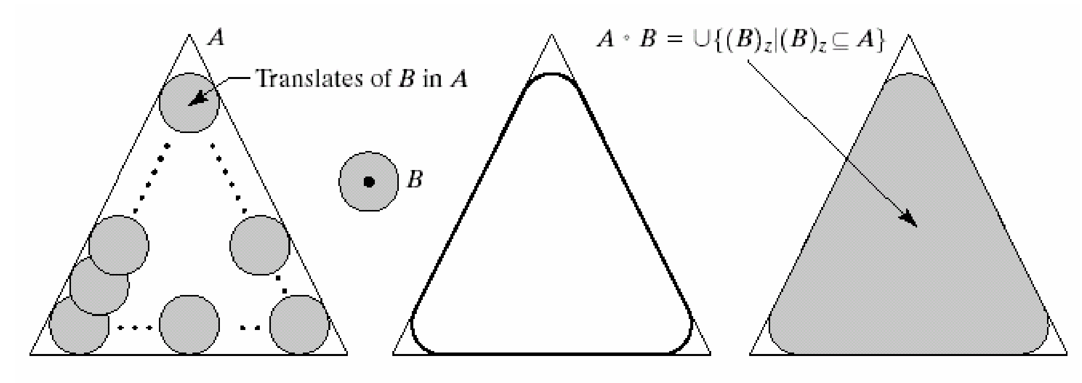


FIGURE 9.8 (a) Structuring element B "rolling" along the inner boundary of A (the dot indicates the origin of B). (c) The heavy line is the outer boundary of the opening. (d) Complete opening (shaded).

闭操作的几何解释



- ✓ $A \bullet B$ 的边界通过B中的点完成
- ✓ B在A的边界外部转动

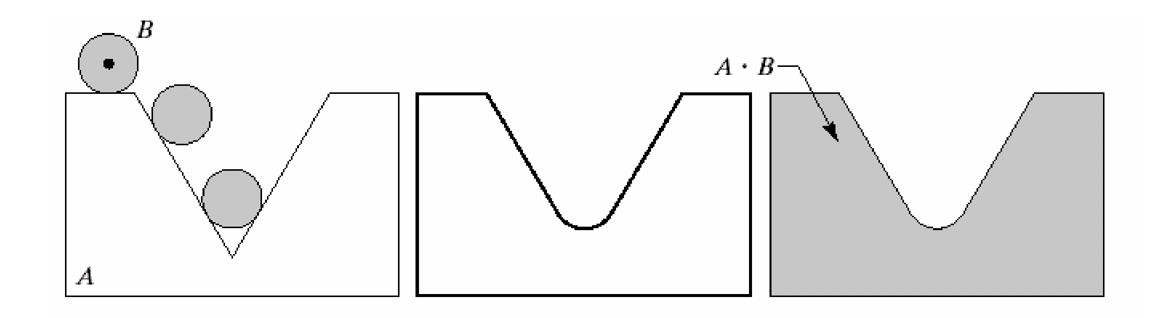


FIGURE 9.9 (a) Structuring element *B* "rolling" on the outer boundary of set *A*. (b) Heavy line is the outer boundary of the closing. (c) Complete closing (shaded).

形态学图像处理



- 开操作的3条性质
  - ✓ $A \circ B$ 是A的子集合
  - ✓ 如果C是D的子集，则 $C \circ B$ 是 $D \circ B$的子集
  - ✓ $(A \circ B) \circ B = A \circ B$
- 闭操作的3条性质
  - ✓ A是$A \bullet B$ 的子集合
  - ✓ 如果C是D的子集，则 $C \bullet B$是$D \bullet B$ 的子集
  - ✓ $(A \bullet B) \bullet B = A \bullet B$

开操作和闭操作应用举例



- 说明：先开操作再闭操作，构成噪声滤波器

  ✓ a图是受噪声污染的指纹二值图像，噪声为黑色背景上的亮元素和亮指纹部分的暗元素

  ✓ b图是使用的结构元素

  ✓ c图是使用结构元素对图a腐蚀的结果：黑色背景噪声消除了，指纹中的噪声尺寸增加

  ✓ d图是使用结构元素对图c膨胀的结果：包含于指纹中的噪声分量的尺寸被减小或被完全消除，带来的问题是：在指纹纹路间产生了新的间断

开操作和闭操作应用举例



- 说明（续）
  - ✓ e图是对图d膨胀的结果，图d的大部分间断被恢复，但指纹的线路变粗了
  - ✓ f图是对图e腐蚀的结果，即对图d中开操作的闭操作。最后结果消除了噪声斑点
  - ✓ 缺点：指纹线路还是有缺点，可以通过加入限制性条件解决

开操作和闭操作应用举例



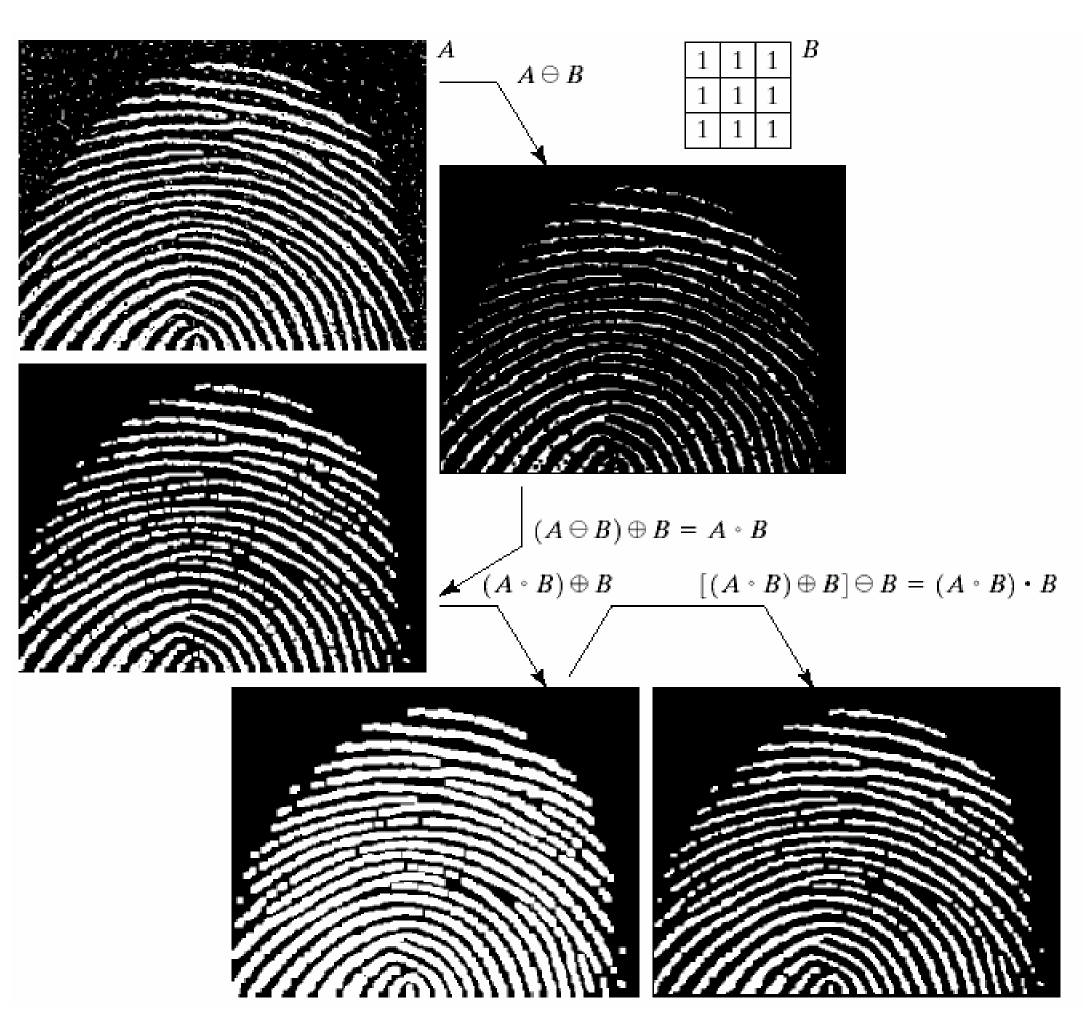
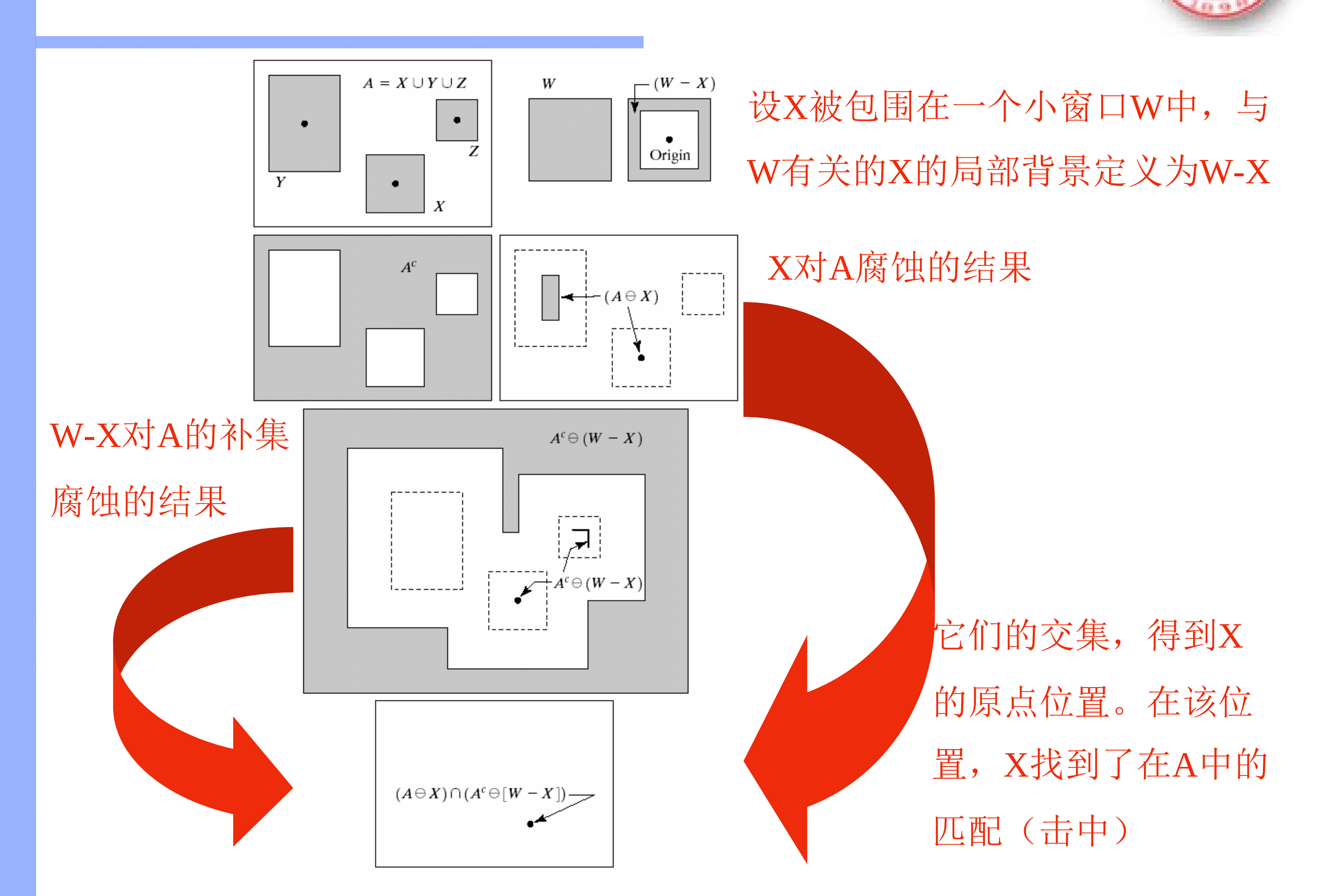




FIGURE 9.11 (a) Noisy image. (c) Eroded image. (d) Opening of A. (d) Dilation of the opening. (e) Closing of the opening. (Original image for this example courtesy of the National Institute of Standards and Technology.)

击中或击不中变换举例——是形状检测的基本工具





击中或击不中变换

- 三个相关公式

A中对B进行的匹配（击中）表示为：

$$A \circledast B = (A \Theta X) \cap [A^c \Theta (W - X)]$$

$$A \circledast B = (A \Theta B_1) \cap (A^c \Theta B_2) \qquad B = (B_1, B_2), B_1 = X, B_2 = (W - X)$$

B_1是由与一个对象相联系的B元素构成的集合，

B_2是与相应背景有关的B元素的集合

$$A \circledast B = (A \Theta B_1) - [A \Theta \hat{B}_2]$$

形态学的主要应用



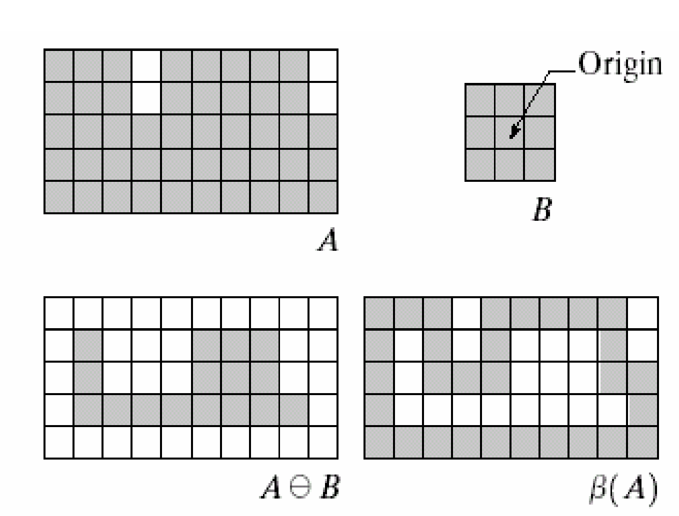
- 边界提取
  - ✓ 边界提取定义为

$$\beta(A) = A - (A \Theta B)$$

  - ✓ 上式表示：先用B对A腐蚀，然后用A减去腐蚀得到，B是结构元素

a	b
c	d

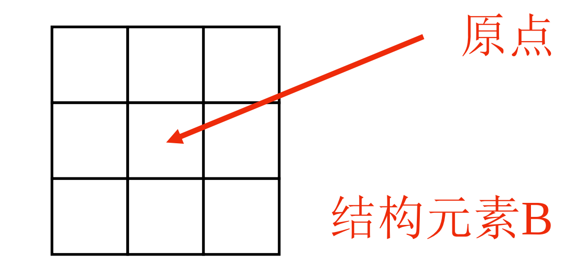
FIGURE 9.13 (a) Set *A*. (b) Structuring element *B*. (c) *A* eroded by *B*. (d) Boundary, given by the set difference between *A* and its erosion.



边界提取举例



- ✓ 1表示为白色，0表示为黑色





a b

FIGURE 9.14 (a) A simple binary image, with 1's represented in white. (b) Result of using Eq. (9.5-1) with the structuring element in Fig. 9.13(b).

形态学的主要应用



- 区域填充

  ✓ 区域填充定义为

$$X_k = (X_{k-1} \oplus B) \cap A^c \quad k = 1,2,3,\ldots$$

X_0=p，如果X_k=X_k-1，则算法在迭代的第k步结束。X_k和A的并集包含被填充的集合和它的边界

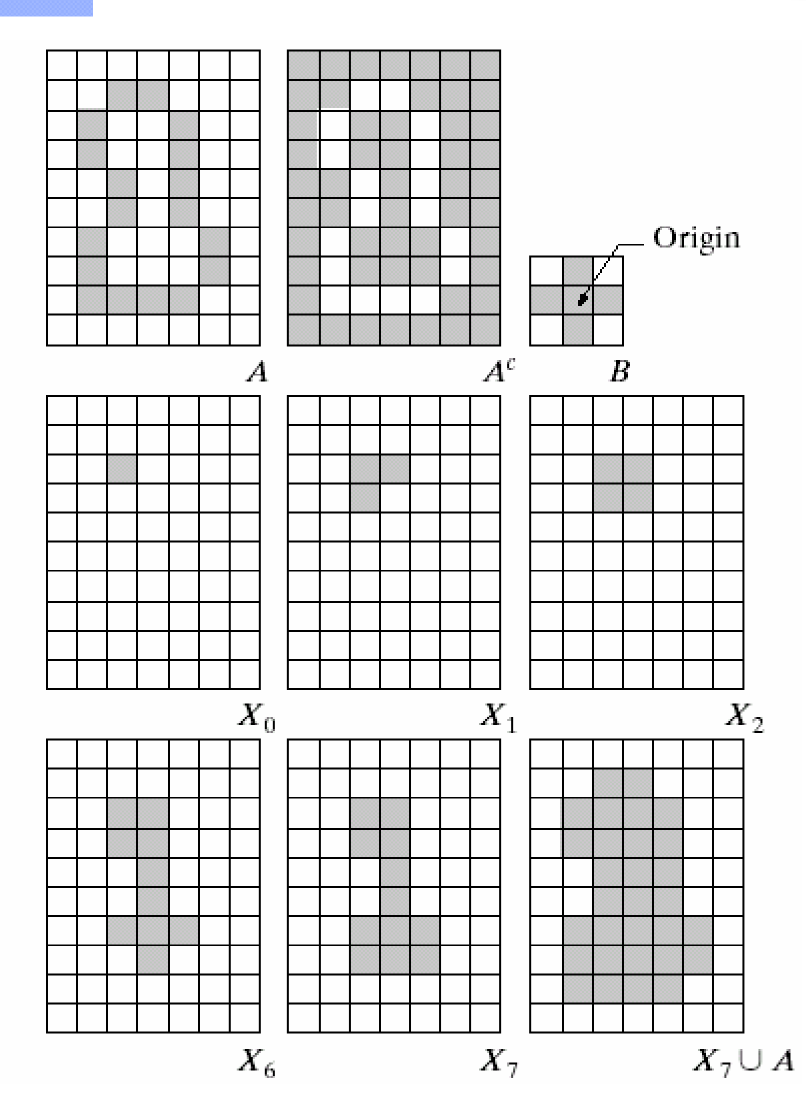
  ✓ 条件膨胀：如果对上述公式的左部不加限制，则上述公式的膨胀将填充整个区域。利用A^c的交集将结果限制在感兴趣区域内，实现条件膨胀

区域填充说明



a	b	c
d	e	f
g	h	i

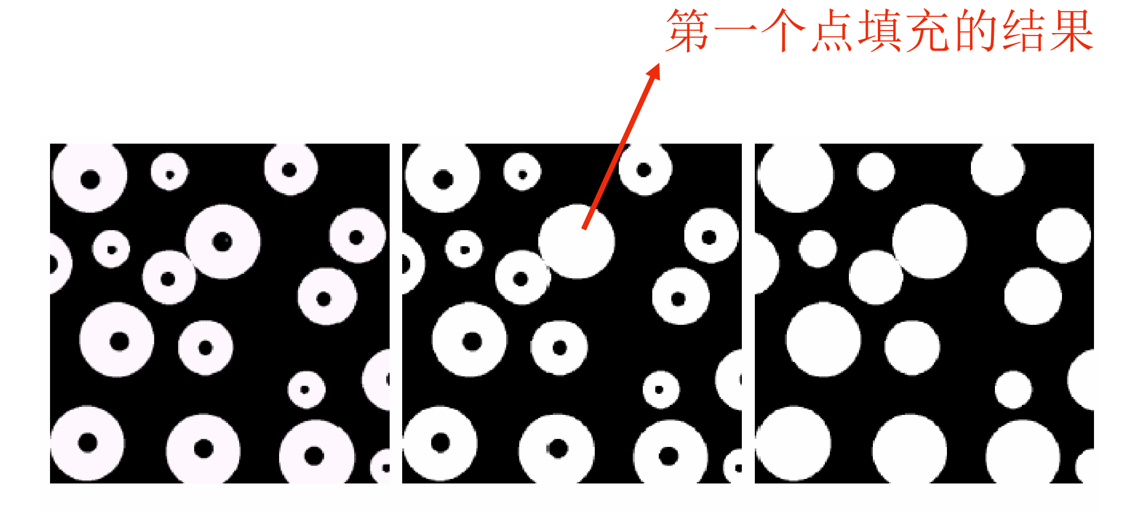
FIGURE 9.15 Region filling. (a) Set A. (b) Complement of A. (c) Structuring element B. (d) Initial point inside the boundary. (e)–(h) Various steps of Eq. (9.5-2). (i) Final result [union of (a) and (h)].





区域填充举例

- ✓ 通过区域填充消除白色圆圈内的黑点



a b c

FIGURE 9.16 (a) Binary image (the white dot inside one of the regions is the starting point for the region-filling algorithm). (b) Result of filling that region (c) Result of filling all regions.

形态学的主要应用



- 连通分量的提取

  ✓ 令Y表示一个包含于集合A中的连通分量，并假设Y中的一个点p是已知的。用下列迭代式生成Y的所有元素：

$$X_k = (X_{k-1} \oplus B) \cap A \quad k = 1,2,3,\ldots$$

x_0=p, 如果X_k=X_k−1，算法收敛，令Y=X_k

连通分量提取说明



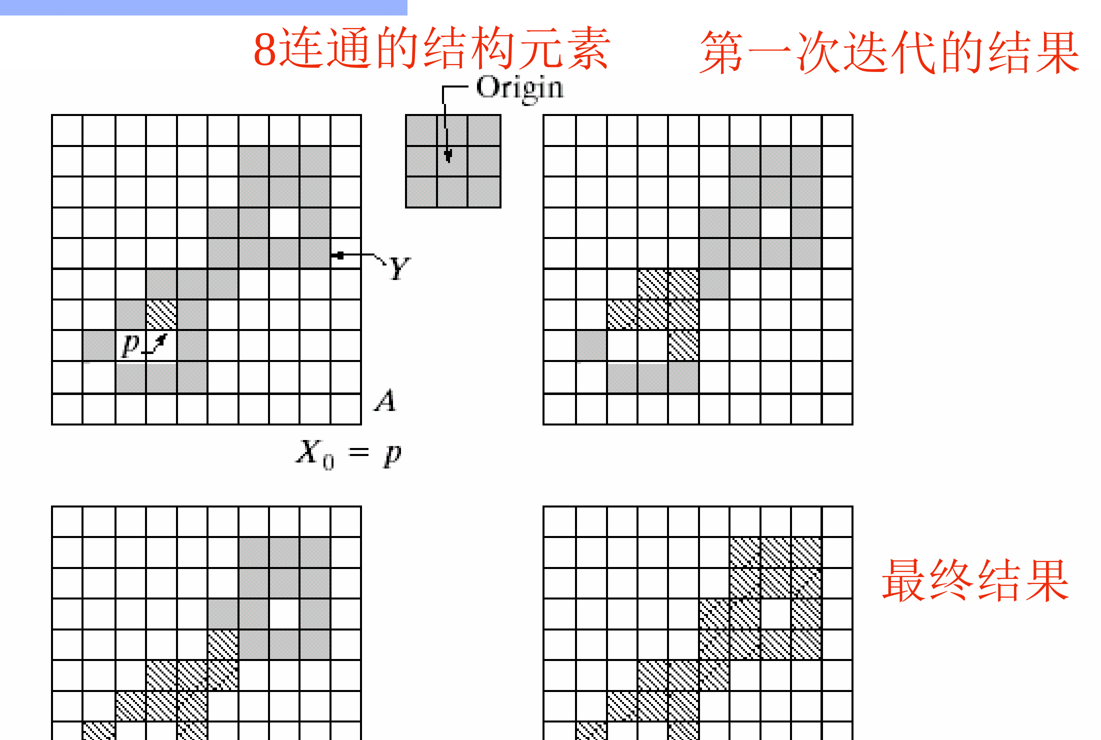


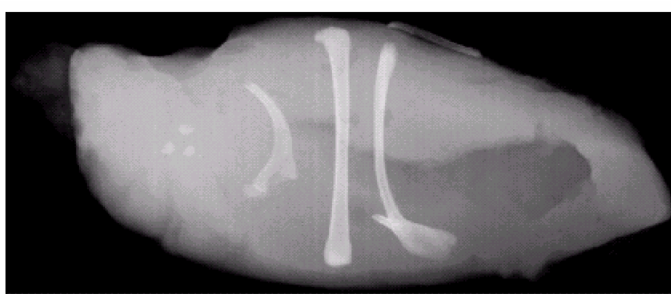
FIGURE 9.17 (a) Set A showing initial point p (all shaded points are valued 1, but are shown different from p to indicate that they have not yet been found by the algorithm). (b) Structuring element. (c) Result of first iterative step. (d) Result of second step. (e) Final result.

连通分量提取的应用举例



a
b
c d

FIGURE 9.18
(a) X-ray image of chicken filet with bone fragments.
(b) Thresholded image. (c) Image eroded with a 5 × 5 structuring element of 1's.
(d) Number of pixels in the connected components of (c). (Image courtesy of NTB Elektronische Geraete GmbH, Diepholz, Germany, www.ntbxray.com.)



含有碎骨的鸡胸X光图像



使用阈值将骨头从背景中提取出来



Connected component	No. of pixels in connected comp
01	11
02	9
03	9
04	39
05	133
06	1
07	1
08	743
09	7
10	11
11	11
12	9
13	9
14	674
15	85

消除细节，对阈值处理后的图像进行腐蚀，保留大尺寸物体

提取连通分量，识别大尺寸对象，其中4个具有最大尺寸

形态学的主要应用



- 凸壳
  - ✓ 如果连接集合A内任意两个点的直线段都在A的内部，则A是凸形的
  - ✓ 集合S的凸壳H是包含S的最小凸集合
  - ✓ H-S称为S的凸缺
  - ✓ 求取集合A的凸壳C(A)的简单形态学算法：

令B^i表示4个结构元素，i=1, 2, 3, 4,

$$X_k^i = \left(X_{k-1} \circledast B^i\right) \cup A \quad i = 1,2,3,4 \qquad k = 1,2,3,4,\ldots$$

$X_0^i = A$ 令 $D^i = X_{conv}^i$，conv表示在$X_k^i = X_{k-1}^i$收敛，则

形态学的主要应用



- 凸壳
  - ✓ $$C(A)=\bigcup_{i=1}^{4} D^{i}$$
  - ✓ 先对A用B^1运用击中或击不中变换，反复使用，当不再发生变化时，执行与A的并集运算，用D^1表示结果
  - ✓ 上述过程用B^2重复，直到不发生变化。
  - ✓ 最后得到的4个D的并集组成了A的凸壳

计算凸壳举例



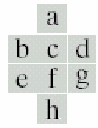
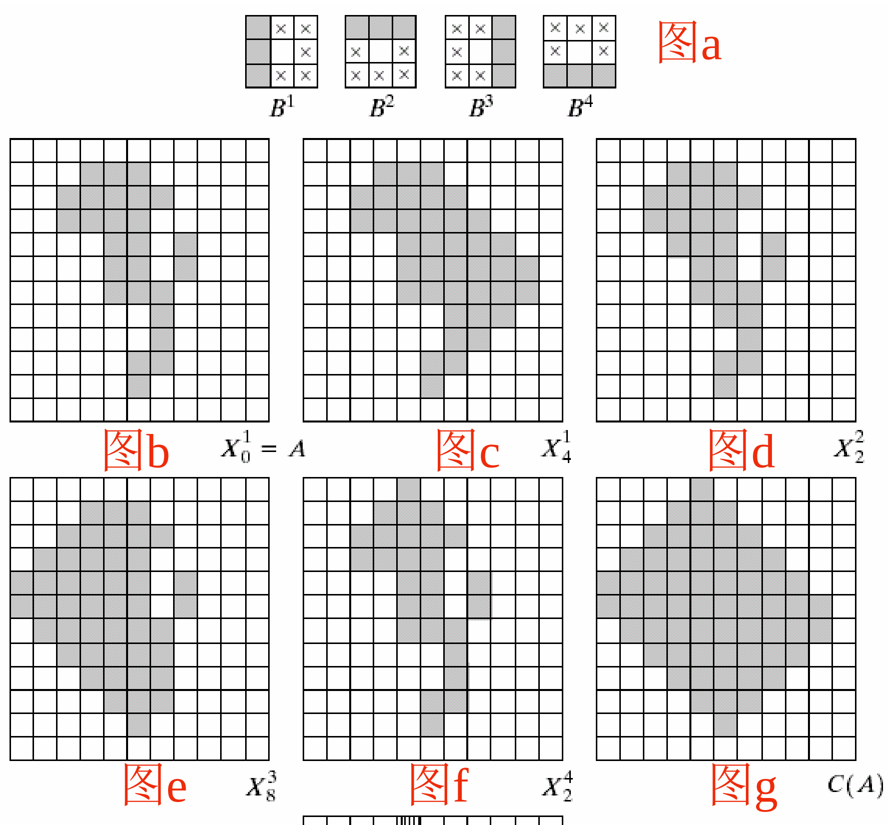
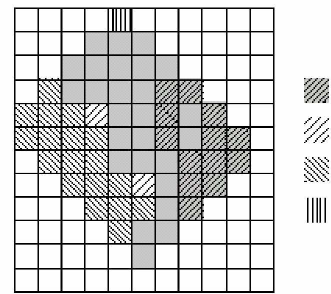


FIGURE 9.19
(a) Structuring elements. (b) Set A. (c)–(f) Results of convergence with the structuring elements shown in (a). (g) Convex hull. (h) Convex hull showing the contribution of each structuring element.





图c-f是用图a中的结构元素得到的收敛结果

图h显示4个结构元素的属性

计算凸壳举例（续）



- 上述过程的一个明显缺点是：凸壳可能超出确保凸性所需的最小尺寸
- 解决办法：限制水平和垂直方向上的尺寸大小，如下图所示。
- 也可限制水平、垂直和对角线方向上的最大尺寸。缺点是增加了算法的复杂性

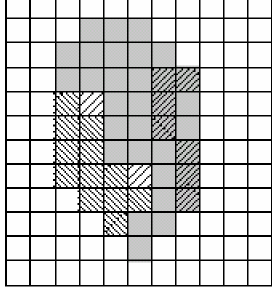


FIGURE 9.20 Result of limiting growth of convex hull algorithm to the maximum dimensions of the original set of points along the vertical and horizontal directions.

形态学的主要应用



- 细化
  - ✓ 细化过程根据击中或击不中变换定义

$$A \otimes B = A - (A \circledast B) = A \cap (A \circledast B)^c$$

  - ✓ 定义结构元素序列为

$$\{B\} = \{B^1, B^2, B^3, \ldots, B^n\}$$

B^i是B^{i-1}旋转后的形式，如在B^4中旋转90^0

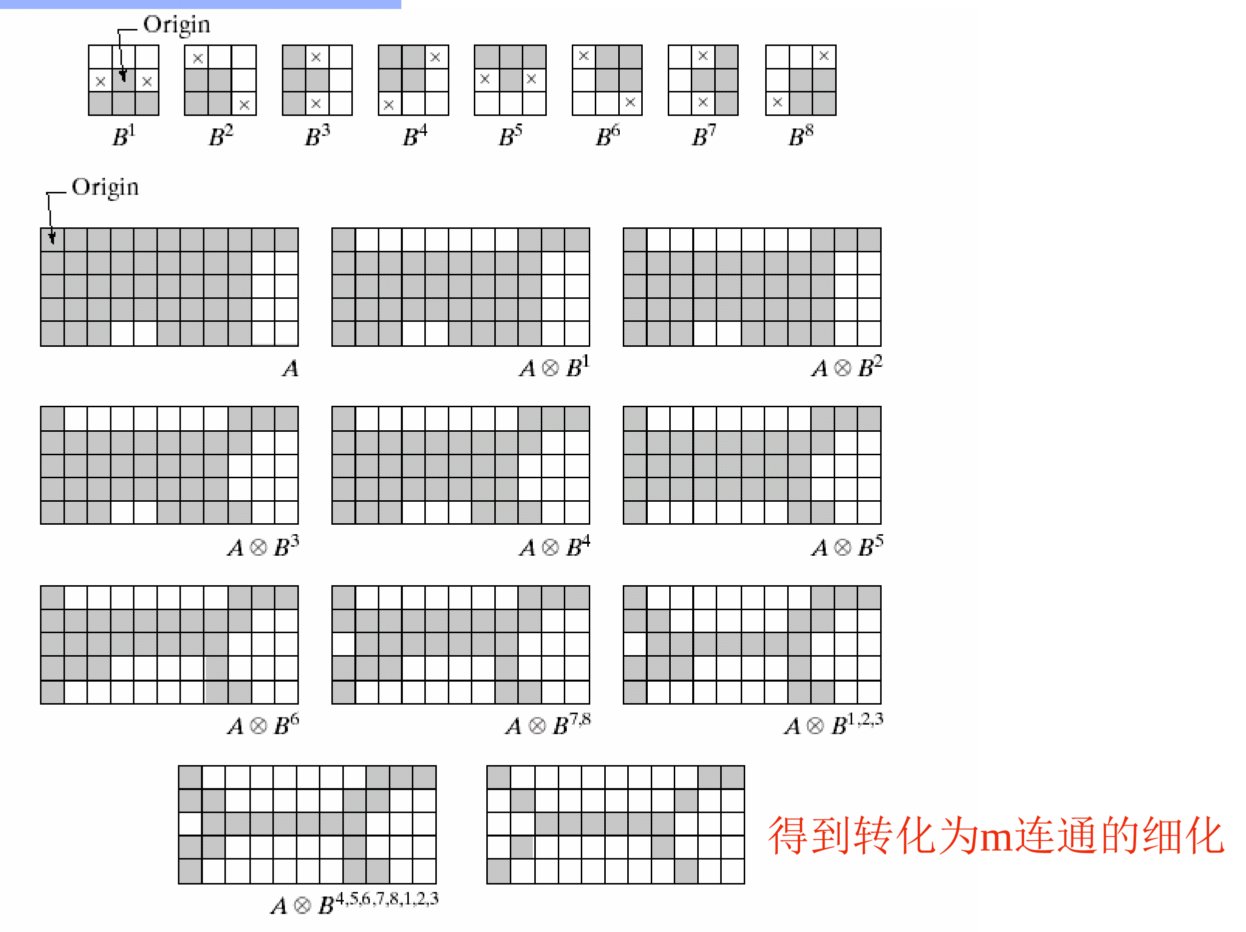
  - ✓ 用结构元素序列定义细化为

$$A \otimes \{B\} = ((\ldots((A \otimes B^1) \otimes B^2) \ldots) \otimes B^n)$$

即连续使用B^1，B^2，…，B^n对A细化

细化过程举例





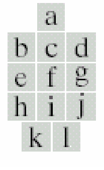


FIGURE 9.21 (a) Sequence of rotated structuring elements used for thinning. (b) Set A. (c) Result of thinning with the first element. (d)–(i) Results of thinning with the next seven elements (there was no change between the seventh and eighth elements). (j) Result of using the first element again (there were no changes for the next two elements). (k) Result after convergence. (l) Conversion to m-connectivity.

形态学的主要应用



- 粗化
  - ✓ 粗化和细化在形态学上是对偶过程，定义为

$$A \odot B = A \cup (A \otimes B)$$

  - ✓ 用结构元素序列定义粗化为

$$A \odot \{B\} = ((\ldots((A \odot B^1) \odot B^2)\ldots) \odot B^n)$$

即连续使用B^1，B^2，…，B^n对A粗化

  - ✓ 粗化可以通过细化算法求补集实现：

先对所讨论集合的背景进行细化，然后对结果求补集，即

形态学的主要应用



- 粗化
  - ✓ 粗化可以通过细化算法求补集实现：

    为了对集合A进行粗化，先令$C=A^c$，然后对C进行细化，最后形成C^c

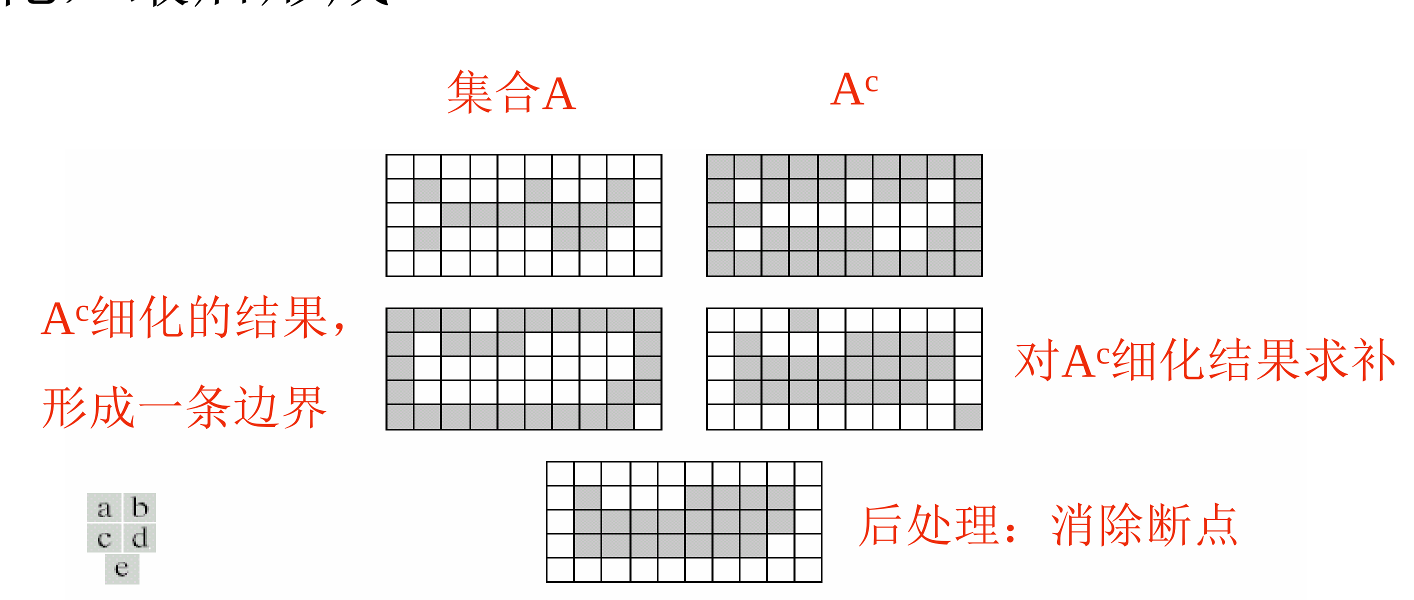


FIGURE 9.22 (a) Set A. (b) Complement of A. (c) Result of thinning the complement of A. (d) Thickened set obtained by complementing (c). (e) Final result, with no disconnected points.



任何问题？

研究生课程



数字图像处理
Digital Image Processing

彭宇新

北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

图像分割



- 概述
- 间断检测
- 边缘连接和边界检测
- 阈值处理
- 基于区域的分割
- 分割中运动的应用

图像分割



- 分割的目的：将图像划分为不同区域
- 三大类方法

✓ 根据区域间灰度不连续搜寻区域之间的边界，在间断检测、边缘连接和边界检测介绍

✓ 以像素性质的分布进行阈值处理，在阈值处理介绍

✓ 直接搜寻区域进行分割，在基于区域的分割中介绍

图像分割



- 概述
  - ✓ 在对图像的研究和应用中，人们往往仅对图像中的某些部分感兴趣，这些部分一般称为目标或前景
  - ✓ 为了辨识和分析目标，需要将有关区域分离提取出来，在此基础上对目标进一步利用，如进行特征提取和测量
  - ✓ 图像分割就是指把图像分成各具特性的区域并提取出感兴趣目标的技术和过程

图像分割

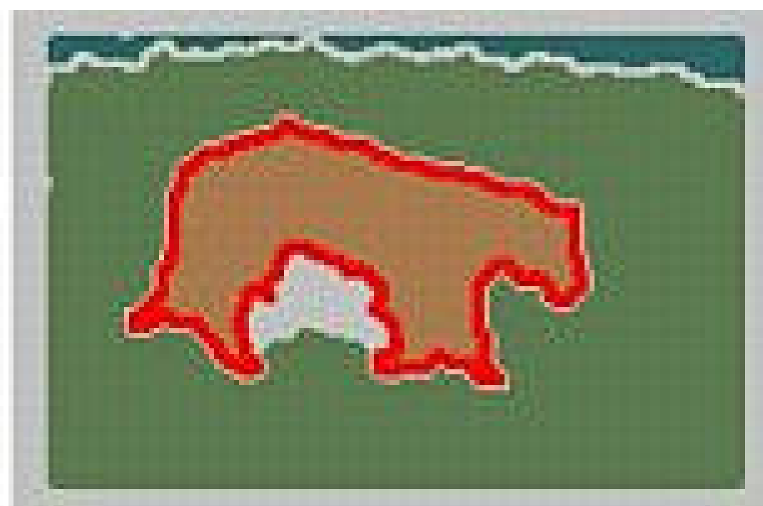


- 概述（续）
  - ✓ 特性可以是灰度、颜色、纹理等，目标可以对应单个区域，也可以对应多个区域
  - ✓ 图像分割算法是基于亮度值的不连续性和相似性
    - ➢ 不连续性是基于亮度的不连续变化分割图像，如图像的边缘
    - ➢ 根据制定的准则将图像分割为相似的区域，如阈值处理、区域生长、区域分离和聚合

图像分割举例







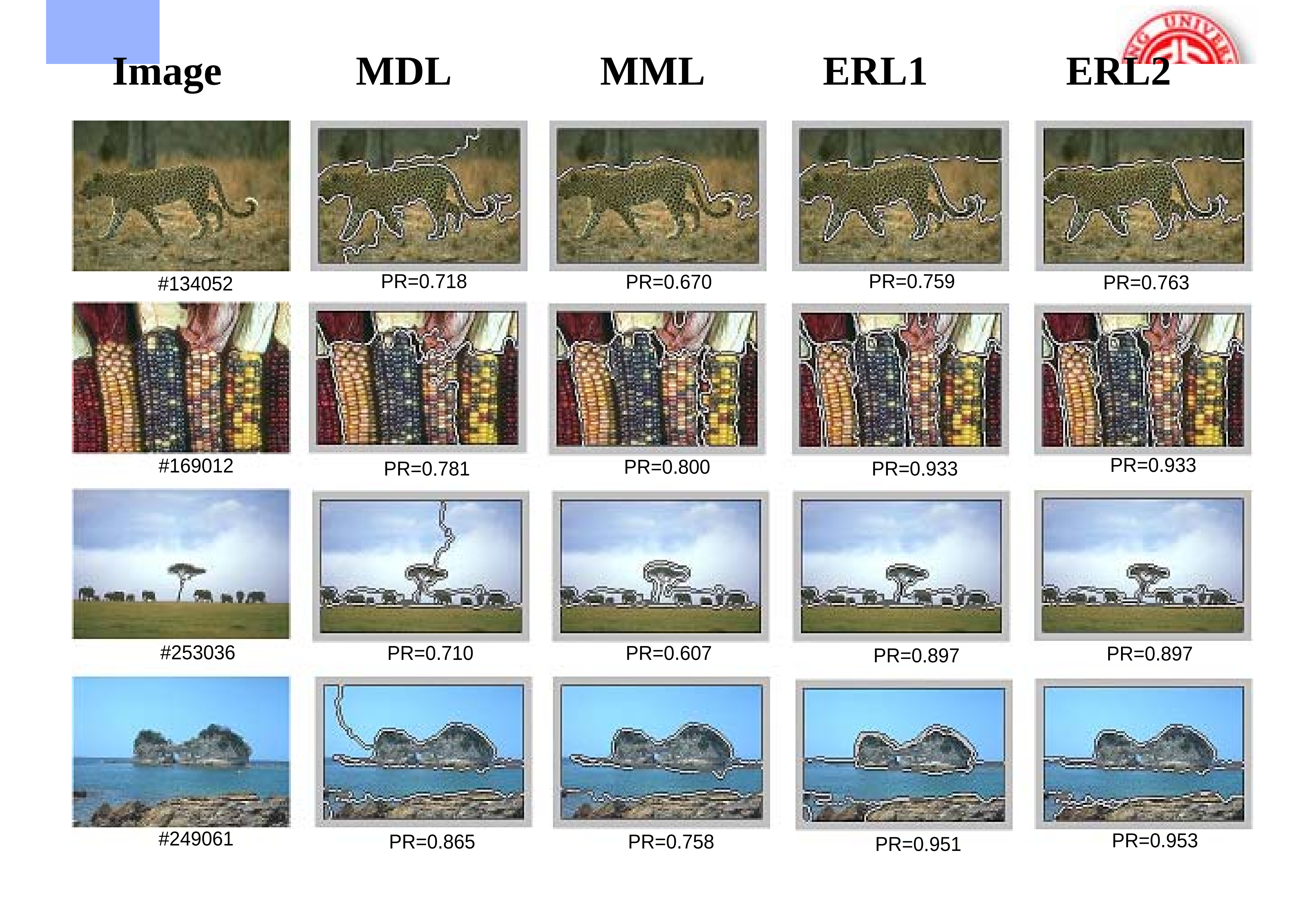


Image
MDL
MML
ERL1
ERL2
#134052
PR=0.718
PR=0.670
PR=0.759
PR=0.763
#169012
PR=0.781
PR=0.800
PR=0.933
PR=0.933
#253036
PR=0.710
PR=0.607
PR=0.897
PR=0.897
#249061
PR=0.865
PR=0.758
PR=0.951
PR=0.953



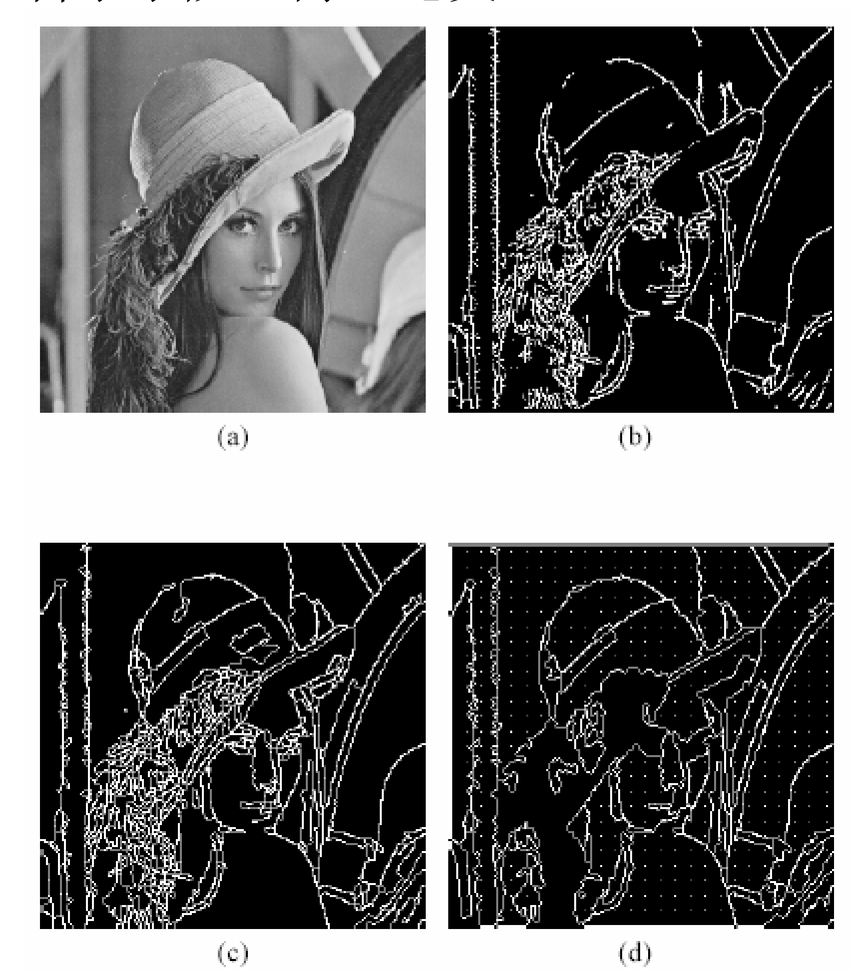
Row 1: Image   Row 2: RPCL   Row 3: CAC   Row 4: ERL

#3096	#85048	#175043	#182053	#219090
pr=0.521	pr=0.740	pr=0.375	pr=0.613	pr=0.822
pr=0.480	pr=0.565	pr=0.401	pr=0.858	pr=0.820
pr=0.861	pr=0.850	pr=0.789	pr=0.890	pr=0.914

基于边缘生长的图像分割算法结果



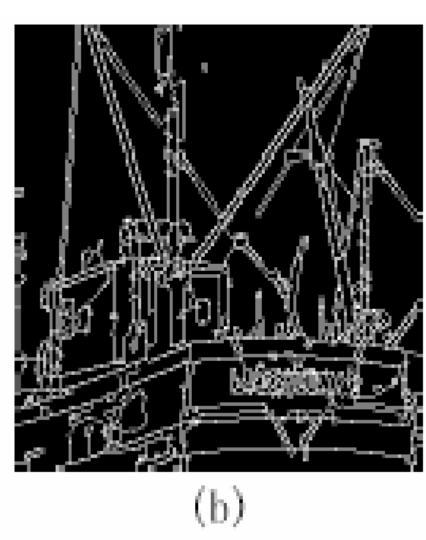
- 参考文献：林通，“基于内容的视频索引与检索方法的研究”，北京大学数学科学学院，博士论文，2001。



基于边缘生长的图像分割算法结果




(a)


(b)


(c)





(a) (b) (c)

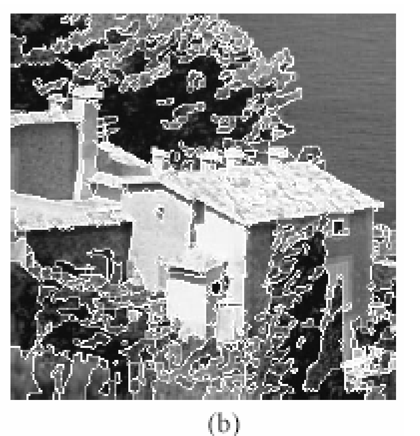
(d) (e) (f)

基于边缘生长的图像分割算法结果



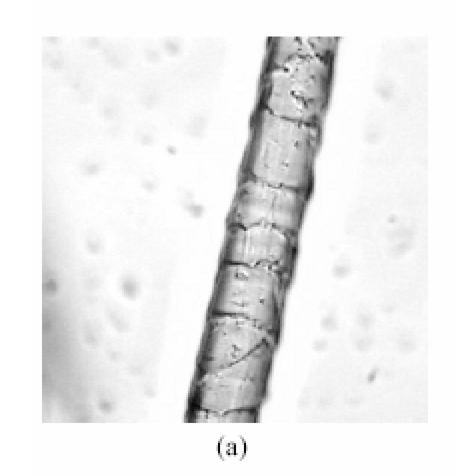


(a)

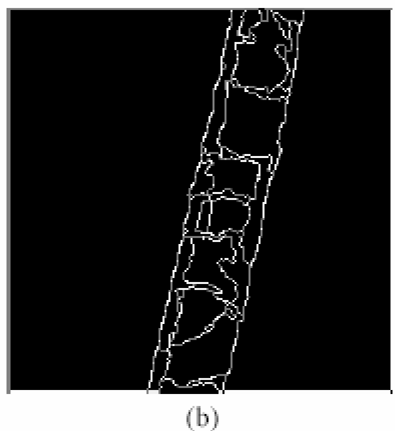


(b)

图 4.8：*house* 图象的分割结果。



(a)

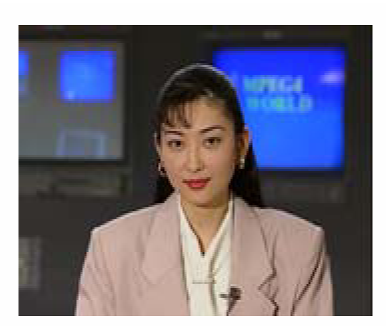


(b)

图 4.9: *wool* 图象的分割结果。

基于边缘生长的图像分割算法结果





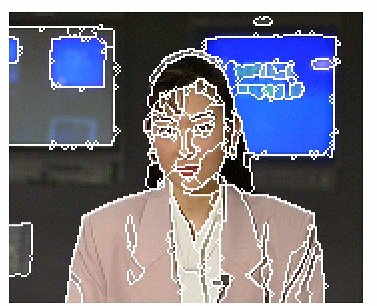


图 4.10: *Akiyo* 视频帧图象的分割结果。

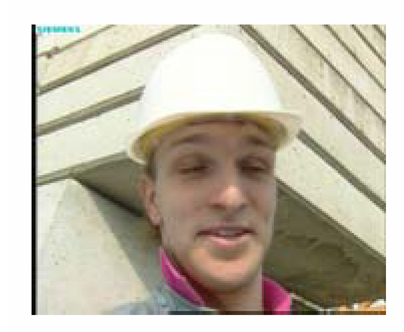




图 4.11: *Foreman* 视频帧图象的分割结果。

图像分割



- 概述
- 间断检测
- 边缘连接和边界检测
- 阈值处理
- 基于区域的分割
- 分割中运动的应用

图像分割



- 间断检测
  - ✓ 点检测
  - ✓ 线检测
  - ✓ 边缘检测
- 寻找间断的一般方法：模板检测

$$R = \omega_1 z_1 + \omega_2 z_2 + \ldots + \omega_9 z_9$$

$$= \sum_{i=1}^{9} \omega_i z_i$$

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

间断检测



- 点检测

使用如图所示的模板，如果 $|R| \geq T$ ，则在模板中心位置检测到一个点

其中，T是阈值，R是模板计算值

-1	-1	-1
-1	8	-1
-1	-1	-1

✓ 基本思想：如果一个孤立点与它周围的点不同，则可以使用上述模板进行检测。

✓ 注意：如果模板响应为0，则表示在灰度级为常数的区域

点检测例子



孔中嵌有一个黑点

应用检测模板的结果

使用图b中像素最高值的90%作为阈值

图a

图b

图c

结论：孤立点可以通过检测模板并设置阈值进行检测

间断检测



- 线检测
  - ✓ 4个线检测模板

FIGURE 10.3 Line masks.

Horizontal

−1	−1	−1
2	2	2
−1	−1	−1

+45°

−1	−1	2
−1	2	−1
2	−1	−1

Vertical

−1	2	−1
−1	2	−1
−1	2	−1

−45°

2	−1	−1
−1	2	−1
−1	−1	2

  - ✓ 第一个模板对水平线有最大响应
  - ✓ 第二个模板对45°方向线有最大响应
  - ✓ 第三个模板对垂直线有最大响应
  - ✓ 第四个模板对-45°方向线有最大响应

间断检测



- 线检测
  - ✓ 用R_1, R_2, R_3和R_4分别代表水平、45°、垂直和-45°方向线的模板响应，在图像中心的点，如果 $|R_i|>|R_j|, j\neq i$

    则此点被认为与在模板i方向上的线更相关
  - ✓ 例：如果$|R_1|>|R_j|, j=2,3,4$

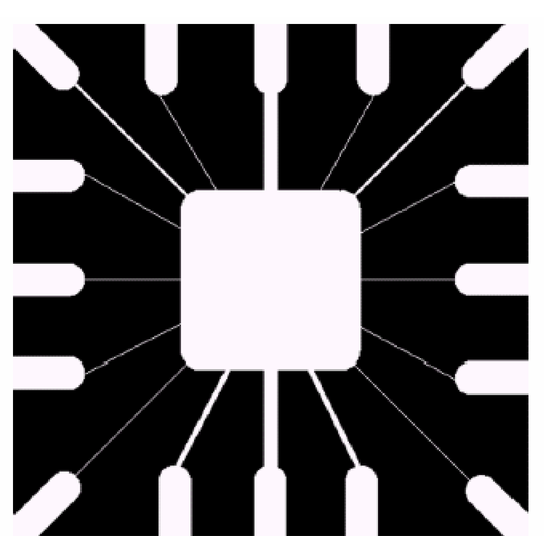
    则该点与水平线有更大的关联
  - ✓ 在灰度恒定的区域，上述4个模板的响应为零

线检测例子



二值电路接线模板，

寻找方向为-45°的线条



a
b c

FIGURE 10.4 Illustration of line detection. (a) Binary wire-bond mask. (b) Absolute value of result after processing with −45° line detector. (c) Result of thresholding image (b).

使用-45°模板，得到结果的绝对值





使用阈值方法，阈值等于图像中最大值得到的结果，一些孤立点使用点模板检测删除，或用形态学腐蚀法删除

间断检测



- 边缘检测
  - ✓ 什么是边缘？

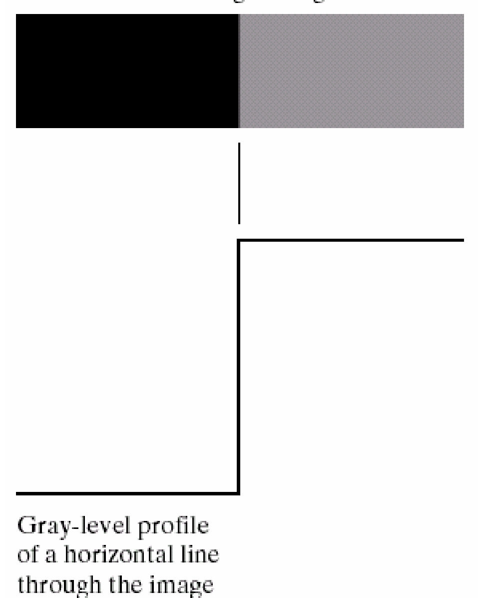
    一组相连的像素集合，这些像素位于两个区域的边界上

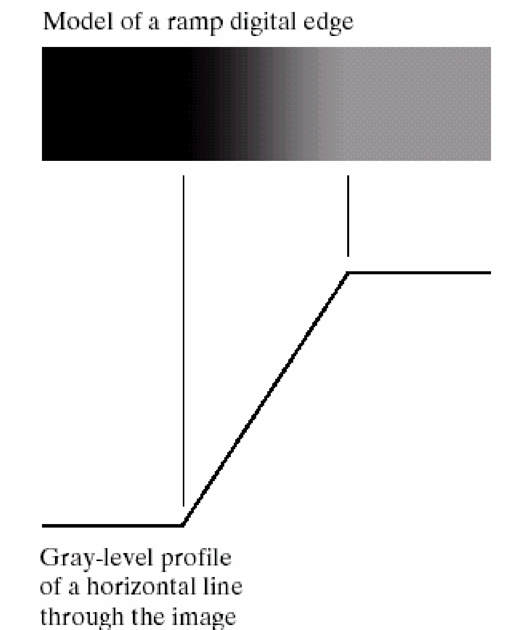
  - ✓ 一阶导数和二阶导数在识别图像边缘中的应用

数字边缘模型



理想数字边缘模型　　　斜坡数字边缘模型





a b

FIGURE 10.5 (a) Model of an ideal digital edge. (b) Model of a ramp edge. The slope of the ramp is proportional to the degree of blurring in the edge.

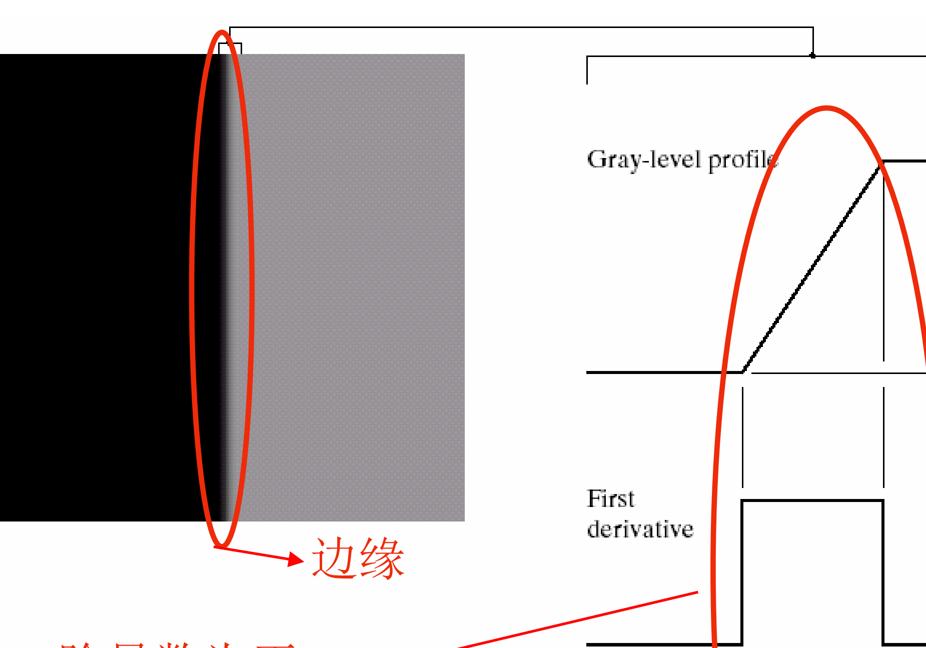
斜坡的产生是由光学系统、取样和图像采集系统的不完善带来的边缘模糊造成的

间断检测





FIGURE 10.6 (a) Two regions separated by a vertical edge. (b) Detail near the edge, showing a gray-level profile, and the first and second derivatives of the profile.



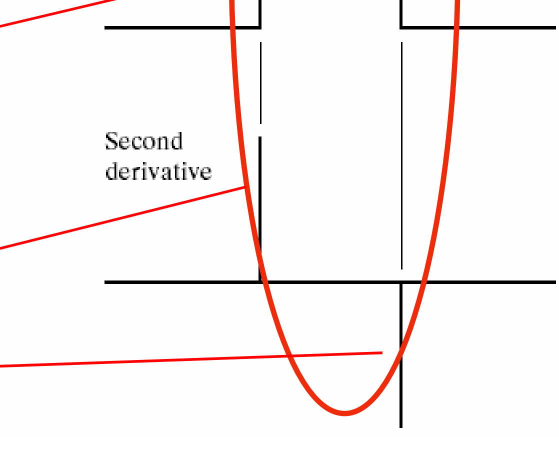
在边缘斜面上，一阶导数为正，

其它区域为零

在边缘与黑色交界处，二阶导数为正

在边缘与亮色交界处，二阶导数为负

沿着斜坡和灰度为常数的区域为零



间断检测



- 结论
  - ✓ 一阶导数可用于检测图像中的一个点是否在边缘上
  - ✓ 二阶导数可以判断一个边缘像素是在边缘亮的一边还是暗的一边
  - ✓ 一条连接二阶导数正值和负值的虚构直线将在边缘中点附近穿过零点
  - ✓ 一阶导数使用梯度算子，二阶导数使用拉普拉斯算子

边缘检测



- 梯度算子
  - ✓ 图像f(x, y)在位置(x, y)的梯度定义为下列向量

$$\nabla F = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

$$\nabla f = mag(\nabla F) = \left[G_x^2 + G_y^2\right]^{\frac{1}{2}}$$

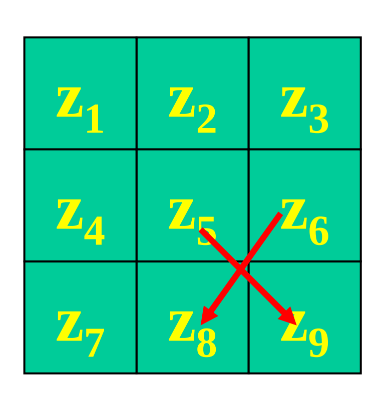
$$\alpha(x, y) = \arctan\left(\frac{G_y}{G_x}\right)$$

边缘检测



- Roberts交叉梯度算子

$$\nabla f \approx |G_x| + |G_y|$$
$$= |z_9 - z_5| + |z_8 - z_6|$$



✓梯度计算由两个模板组成，第一个求得梯度的第一项，第二个求得梯度的第二项，然后求和，得到梯度。

✓两个模板称为Roberts交叉梯度算子

-1	0
0	1

0	-1
1	0

边缘检测



- Prewitt梯度算子——3x3的梯度模板

$$\nabla f \approx |G_x| + |G_y|$$
$$= |(z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)| +$$
$$|(z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)|$$

-1	-1	-1
0	0	0
1	1	1

-1	0	1
-1	0	1
-1	0	1

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

边缘检测



- Sobel梯度算子——3x3的梯度模板

  权值2用于通过增加中心点的重要性而实现某种程度的平滑效果

$$\nabla f \approx |G_x| + |G_y|$$

$$= |(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)| +$$

$$|(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)|$$

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

边缘检测



- 结论
  - ✓ Prewitt和Sobel算子是计算数字梯度时最常用的算子
  - ✓ Prewitt模板比Sobel模板简单，但Sobel模板能够有效抑制噪声



边缘检测

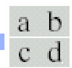


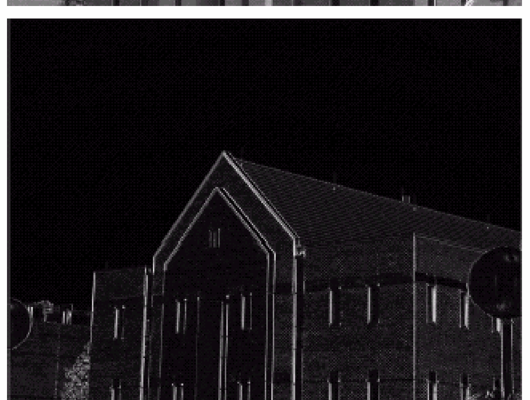
FIGURE 10.10 (a) Original image. (b) $|G_x|$, component of the gradient in the x-direction. (c) $|G_y|$, component in the y-direction. (d) Gradient image, $|G_x| + |G_y|$.

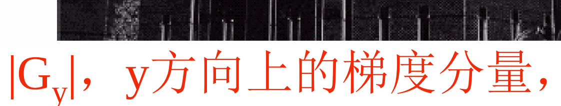
原图像



$|G_x|$，x方向上的梯度分量，水平细节非常清楚







$|G_y|$，y方向上的梯度分量，垂直细节非常清楚

梯度图像$|G_x|+|G_y|$，水平和垂直细节都非常清楚

问题：砖墙等图像细节对边缘提取不必要

处理办法：对图像进行平滑处理

边缘检测



原图像经过5×5的均值滤波器进行平滑处理

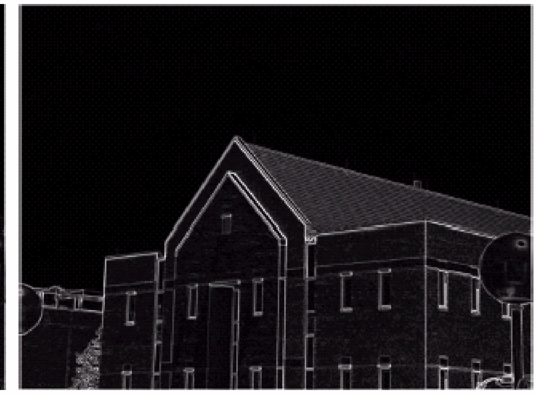
$|G_x|$，x方向上的梯度分量，突出水平细节



a b
c d

FIGURE 10.11 Same sequence as in Fig. 10.10, but with the original image smoothed with a 5 × 5 averaging filter.





$|G_y|$，y方向上的梯度分量，突出垂直细节

梯度图像$|G_x|+|G_y|$，突出水平和垂直细节

边缘检测



Sobel45^{o}模板的检测结果　　Sobel-45°模板的检测结果





a b

FIGURE 10.12 Diagonal edge detection. (a) Result of using the mask in Fig. 10.9(c). (b) Result of using the mask in Fig. 10.9(d). The input in both cases was Fig. 10.11(a).

拉普拉斯算子



- 图像函数的拉普拉斯变换定义为

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

$$\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8)$$

$$\nabla^2 f = 8z_5 - (z_1 + z_2 + z_3 + z_4 + z_6 + z_7 + z_8 + z_9)$$

FIGURE 10.13 Laplacian masks used to implement Eqs. (10.1-14) and (10.1-15), respectively.

0	−1	0
−1	4	−1
0	−1	0

−1	−1	−1
−1	8	−1
−1	−1	−1

拉普拉斯算子



- 拉普拉斯算子总结
  - 缺点：
    - 拉普拉斯算子对噪声具有敏感性
    - 拉普拉斯算子的幅值产生双边缘
    - 拉普拉斯算子不能检测边缘的方向
  - 优点：
    - 可以利用零交叉的性质进行边缘定位
    - 可以确定一个像素是在边缘暗的一边还是亮的一边

拉普拉斯算子



- 拉普拉斯算子与平滑过程一起利用零交叉找到边缘

$$h(r) = -e^{\frac{r^2}{2\sigma^2}} \qquad r^2 = x^2 + y^2$$

σ 是标准差

$$\nabla^2 h(r) = -\left[\frac{r^2 - \sigma^2}{\sigma^4}\right] e^{-\frac{r^2}{2\sigma^2}}$$

上式称为高斯型的拉普拉斯算子LoG

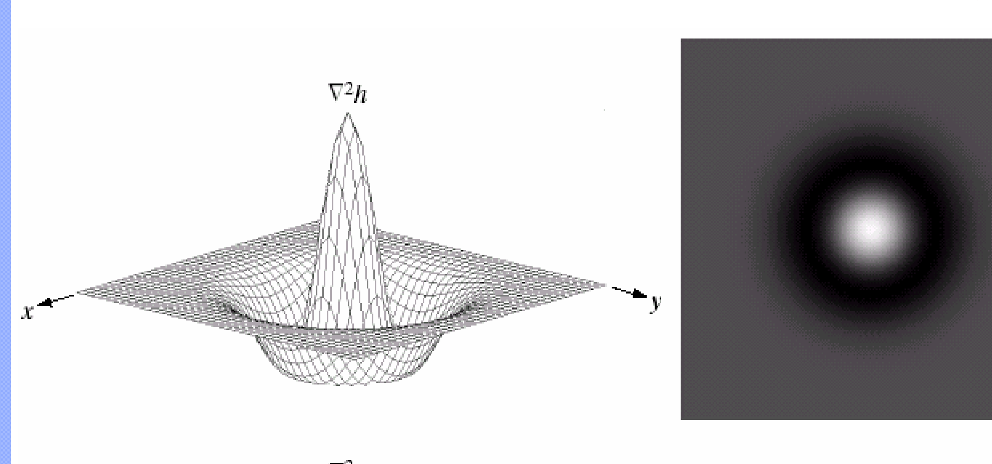
拉普拉斯算子



三维曲线

图像：白色是正值区域，黑色是负值区域，

灰色是零值区域



a	b
c	d

FIGURE 10.14 Laplacian of a Gaussian (LoG). (a) 3-D plot. (b) Image (black is negative, gray is the zero plane, and white is positive). (c) Cross section showing zero crossings. (d) 5 × 5 mask approximation to the shape of (a).

$\nabla^2 h$

r

$-\sigma$ σ

横截面

0	0	−1	0	0
0	−1	−2	−1	0
−1	−2	16	−2	−1
0	−1	−2	−1	0
0	0	−1	0	0

近似的5×5模板：一个正的中心项，周围是一个相邻的负值区域，并被一个零值的外部区域包围。系数的总和为零

拉普拉斯算子

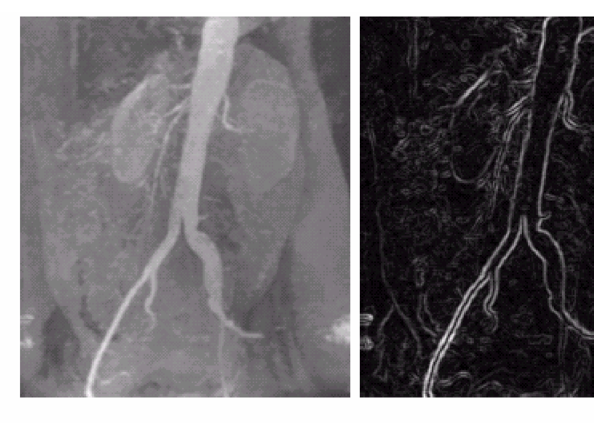


- 高斯型拉普拉斯算子总结
  - ✓ 高斯型函数的目的是对图像进行平滑处理
  - ✓ 拉普拉斯算子的目的是提供一幅用零交叉确定边缘位置的图像
  - ✓ 平滑处理减少了噪声的影响

例：通过零交叉检测边缘

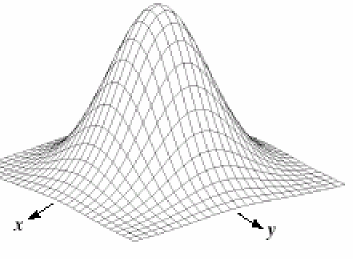






Sobel算子检测结果

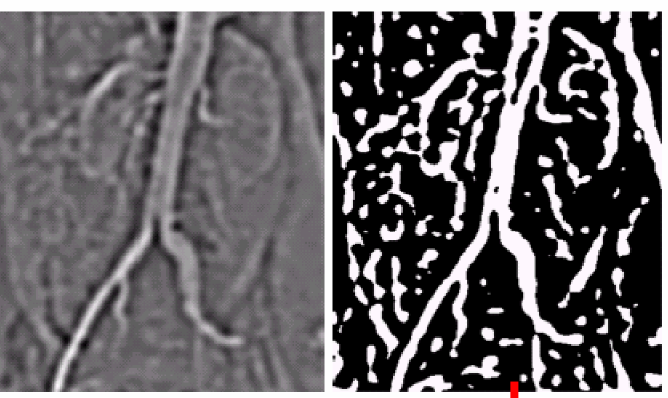
空间高斯型平滑函数

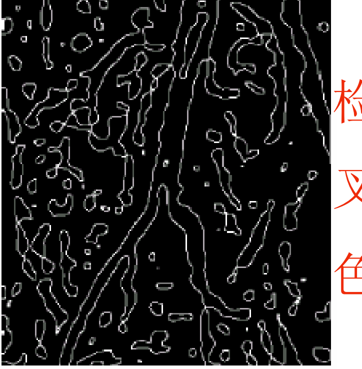


–1	–1	–1
–1	8	–1
–1	–1	–1

拉普拉斯算子模板

LoG检测结果





检测边缘：寻找零交叉点，检测黑色和白色区域之间的过渡点

对LoG图像设置阈值的结果，所有正值区域为白色，所有负值区域为黑色

拉普拉斯算子



- 结论（对比二阶拉普拉斯算子和一阶Sobel梯度算子）
  - ✓ 缺点
    - ➢ 边缘由许多闭合环的零交叉点决定
    - ➢ 零交叉点的计算比较复杂
  - ✓ 优点
    - ➢ 零交叉点图像中的边缘比梯度边缘细
    - ➢ 抑制噪声的能力和反干扰性能
  - ✓ 结论：梯度算子具有更多的应用

图像分割



- 概述
- 间断检测
- 边缘连接和边界检测
- 阈值处理
- 基于区域的分割
- 分割中运动的应用

图像分割



- 边缘连接和边界检测
  - ✓ 为什么需要边缘连接？
  - ✓ 局部处理
  - ✓ 整体处理之霍夫变换

图像分割



- 为什么需要边缘连接？
  - ✓ 由于噪声、照明等产生边缘间断，使得一组像素难以完整形成边缘
  - ✓ 因此，在边缘检测算法后，使用连接过程将间断的边缘像素组合成完整边缘

图像分割



- 局部处理
  - ✓ 分析图像中每个边缘点(x, y)的一个邻域内的像素，根据某种准则将相似点进行连接，由满足该准则的像素连接形成边缘
  - ✓ 如何确定边缘像素的相似性
    - 边缘像素梯度算子的响应强度
    - 边缘像素梯度算子的方向

图像分割



- 边缘像素梯度算子的响应强度

如果 $\left|\nabla f(x,y)-\nabla f(x_0,y_0)\right|\le E$

则 (x, y) 邻域内坐标为 (x_0, y_0) 的边缘像素，在幅度上相似于 (x, y) 的像素

- 边缘像素梯度算子的方向

如果 $\left|\alpha(x,y)-\alpha(x_0,y_0)\right|< A,\quad \alpha(x,y)=\arctan\left(\frac{G_y}{G_x}\right)$

则(x,y)邻域内坐标为(x_0,y_0)的边缘像素，在角度上相似于(x,y)的像素

边缘局部处理例子



原始图像：检测车牌牌照

Sobel算子的G_y分量

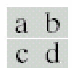
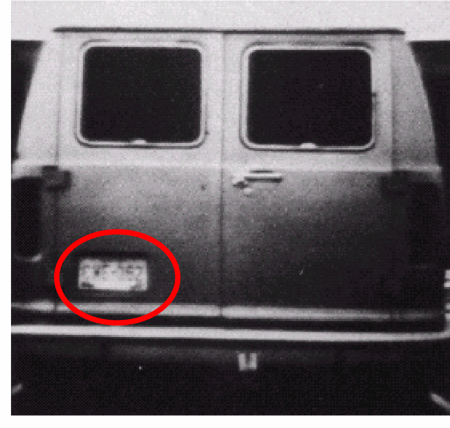
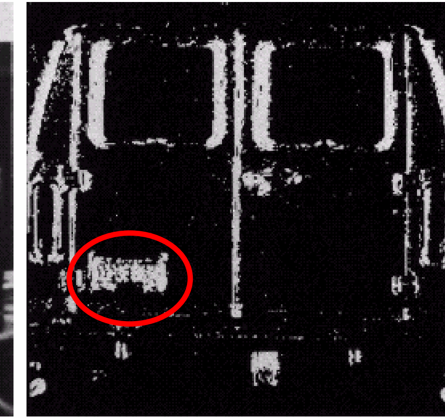
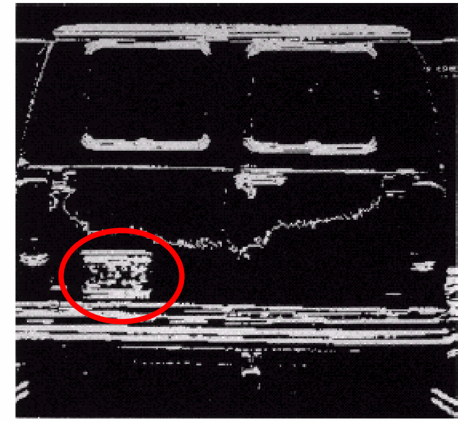
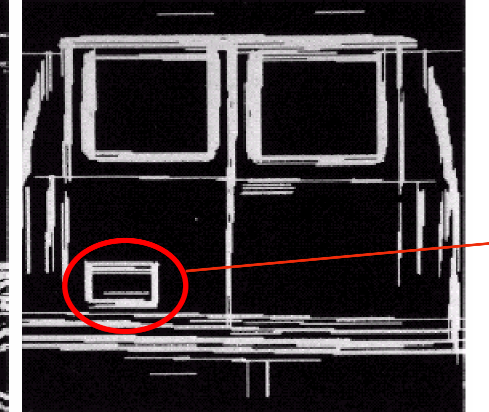


FIGURE 10.16
(a) Input image.
(b) G_y component of the gradient.
(c) G_x component of the gradient.
(d) Result of edge linking. (Courtesy of Perceptics Corporation.)









美国牌照的长宽比例是2:1

Sobel算子的G_x分量

边缘连接的结果

通过Hough变换进行整体处理

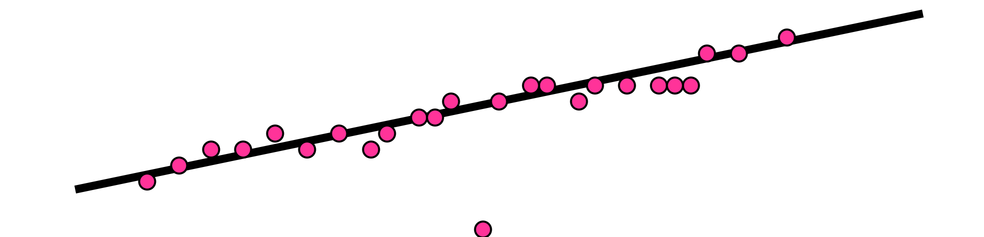


- Hough变换
  - ✓ 问题的提出
  - ✓ Hough变换的基本思想
  - ✓ 算法实现
  - ✓ Hough变换的扩展

通过Hough变换进行整体处理



- Hough变换问题的提出
  - ✓ 在找出边界点集之后，需要连接，形成完整的边界图形描述



通过Hough变换进行整体处理



- Hough变换的基本思想
  - ✓ 对于边界上的n个点的点集，找出共线的点集和直线方程。
  - ✓ 对于任意两点的直线方程：y = ax + b，构造一个参数a，b的平面，从而有如下结论：

通过Hough变换进行整体处理

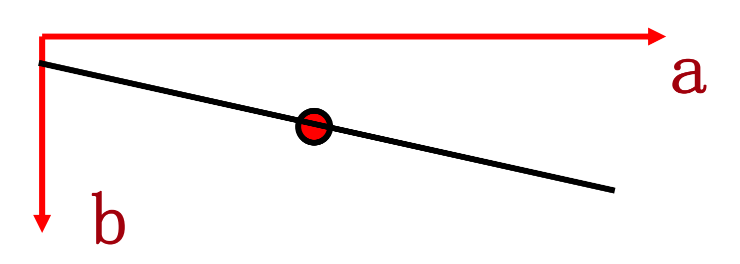


- Hough变换的基本思想
  - ✓ xy平面上的任意一条直线y = ax + b，对应在参数ab平面上都有一个点

a

b

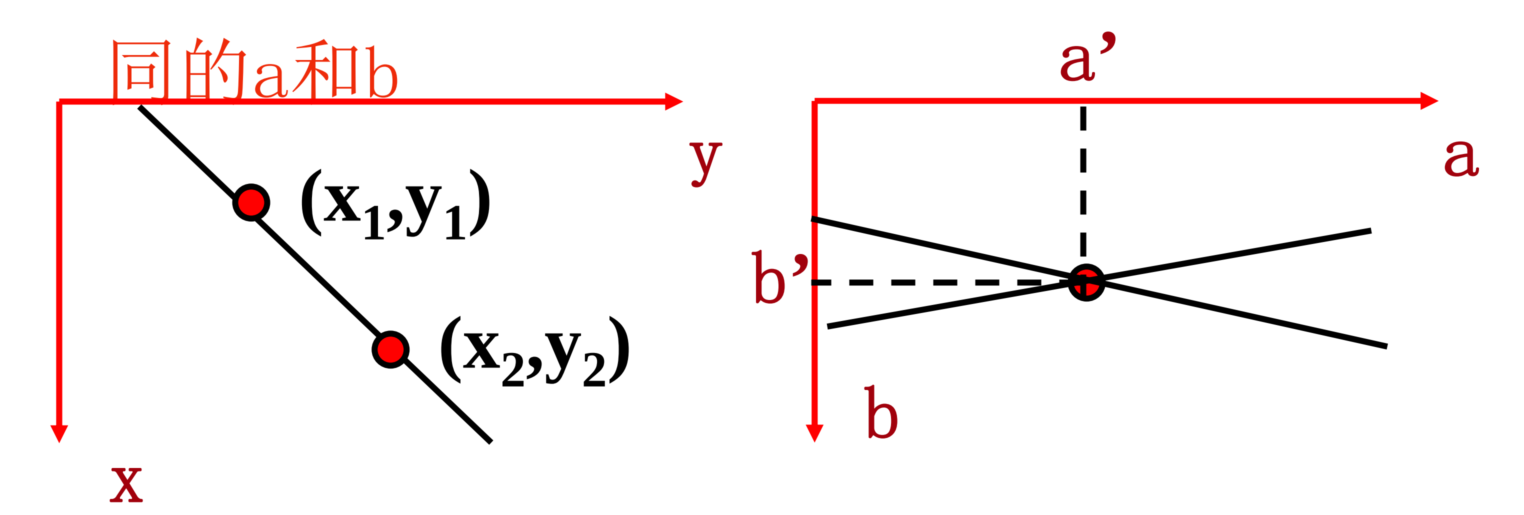
  - ✓ 过xy平面一个点(x, y)的所有直线，构成参数ab平面上的一条直线



通过Hough变换进行整体处理



- Hough变换的基本思想
  - ✓ 如果点(x_1, y_1)与点(x_2, y_2)共线，那么这两点在参数ab平面上的直线将有一个交点,具有相同的a和b

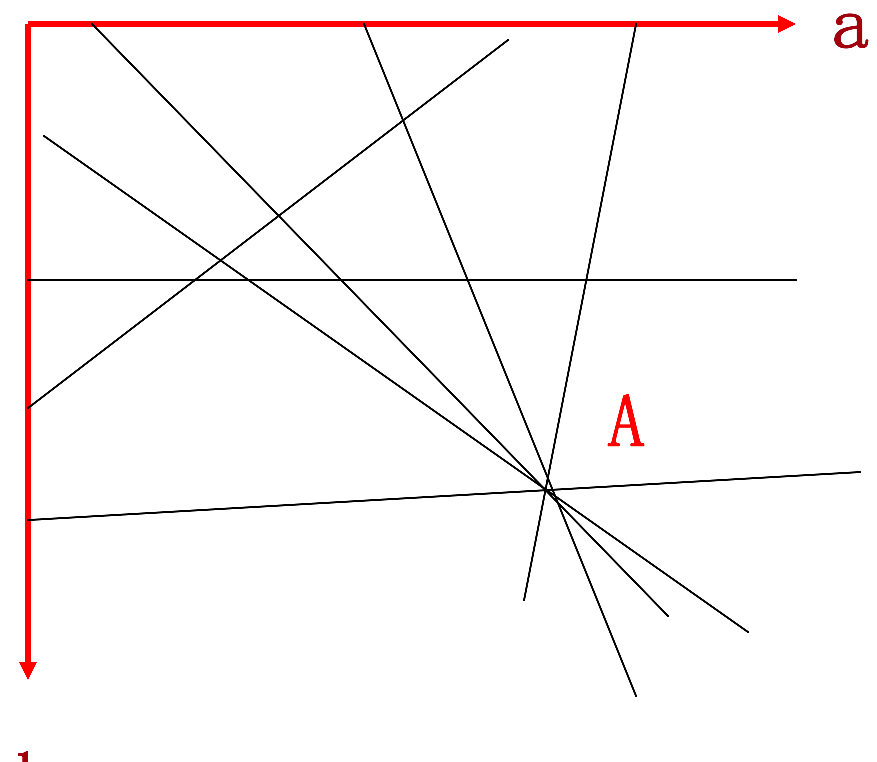


  - ✓ 在参数ab平面上相交直线最多的点，对应的xy平面上的直线就是我们的解





- Hough变换的基本思想



通过Hough变换进行整体处理



- Hough变换算法实现
  - ✓ 由于垂直直线a为无穷大，我们改用极坐标形式：$x\cos\theta + y\sin\theta = \rho$
  - ✓ 参数平面为θ, ρ ，对应不是直线而是正弦曲线
  - ✓ 使用交点累加器，或交点统计直方图，找出相交线段最多的参数空间的点
  - ✓ 然后找出该点对应的xy平面的直线线段

通过Hough变换进行整体处理



- Hough变换的扩展
  - ✓ Hough变换不只对直线，也可以用于圆：

    $$(x - c_1)^2 + (y - c_2)^2 = c_3^2$$

  - ✓ 这时需要三个参数的参数空间

图像分割



- 概述
- 间断检测
- 边缘连接和边界检测
- 阈值处理
- 基于区域的分割
- 分割中运动的应用

图像分割



- 阈值处理
  - ✓ 基础
  - ✓ 基本全局阈值
  - ✓ 基本自适应阈值
  - ✓ 最佳全局和自适应阈值
  - ✓ 通过边界特性选择阈值
  - ✓ 基于不同变量的阈值



阈值处理

- 基础

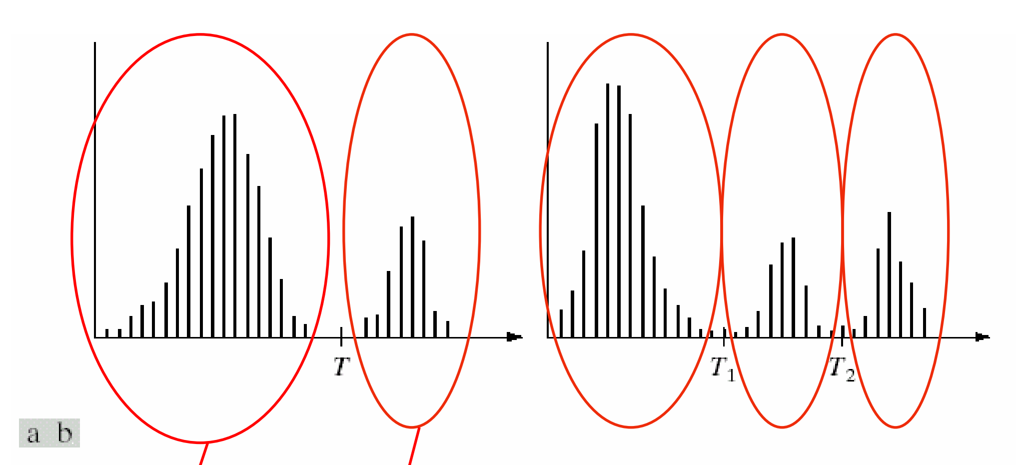


FIGURE 10.26 (a) Gray-level histograms that can be partitioned by (a) a single threshold, and (b) multiple thresholds.

暗的背景：$f(x,y) \leqslant T$

亮的对象：$f(x,y) > T$

暗的背景：$f(x,y) \leqslant T1$

亮的一个对象：　$T_1 < f(x,y) \leqslant T_2$

亮的另一个对象：$f(x,y) > T_2$

阈值处理



- 基础
  - ✓ 阈值处理操作

$$T = T[x, y, p(x, y), f(x, y)]$$

f(x, y)是点(x, y)的灰度级，p(x, y)表示该点的局部性质，如以(x, y)为中心的邻域的平均灰度级

  - ✓ 阈值处理后的图像g(x, y)定义为

$$g(x,y)=\begin{cases}1 & f(x,y)>T \\ 0 & f(x,y)\le T\end{cases}$$

阈值处理



- 基础

$$g(x,y)=\begin{cases}1 & f(x,y)>T\\0 & f(x,y)\le T\end{cases}$$

✓ 标记为1的像素对应于对象，标记为0的像素对应于背景

✓ 当T仅取决于f(x, y)，阈值称为全局的

✓ 当T取决于f(x, y)和p(x, y)，阈值是局部的

✓ 当T取决于空间坐标x和y，阈值就是动态的或自适应的

图像分割



- 基本全局阈值例子

亮背景

暗对象

0 63 127 191 255

a
b c

FIGURE 10.28 (a) Original image. (b) Image histogram. (c) Result of global thresholding with *T* midway between the maximum and minimum gray levels.

图像分割

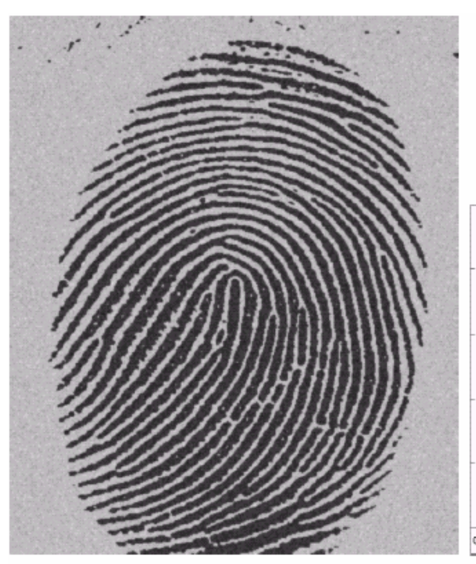


- 计算基本全局阈值算法
  1. 选择一个T的初始估计值
  2. 用T分割图像，生成两组像素：G_1由所有灰度值大于T的像素组成，而G_2由所有灰度值小于或等于T的像素组成
  3. 对区域G_1和G_2中的所有像素计算平均灰度值μ_1和μ_2
  4. 计算新的阈值 $T = \frac{1}{2}(\mu_1 + \mu_2)$
  5. 重复步骤2到4，直到逐次迭代所得的T值之差小于事先定义的参数T_0

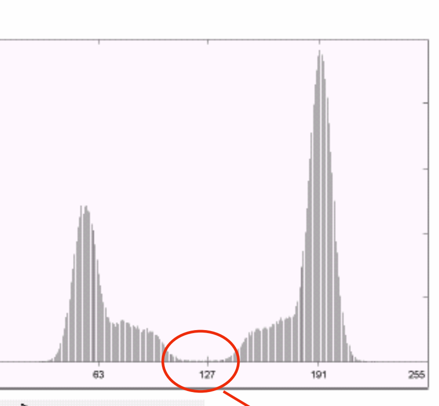
利用基本全局阈值算法的例子



原图



原图的直方图



a b
c

FIGURE 10.29 (a) Original image. (b) Image histogram. (c) Result of segmentation with the threshold estimated by iteration. (Original courtesy of the National Institute of Standards and Technology.)

波谷作为阈值

基本全局阈值算法处理的结果

T0=0，3次迭代得到值为125.4

最后确定T=125



图像分割



- 基本自适应阈值
  - ✓ 单一全局阈值存在的问题
    - ➢ 不均匀亮度图像无法有效分割
  - ✓ 方法
    - ➢ 将图像进一步细分为子图像，并对不同的子图像使用不同的阈值处理
    - ➢ 解决的关键问题：如何将图像进行细分和如何为得到的子图像估计阈值
    - ➢ 自适应阈值：取决于像素在子图像中的位置

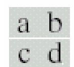
基本自适应阈值举例



一个全局阈值处理后的结果：

人工设置直方图的波谷为阈值

原图



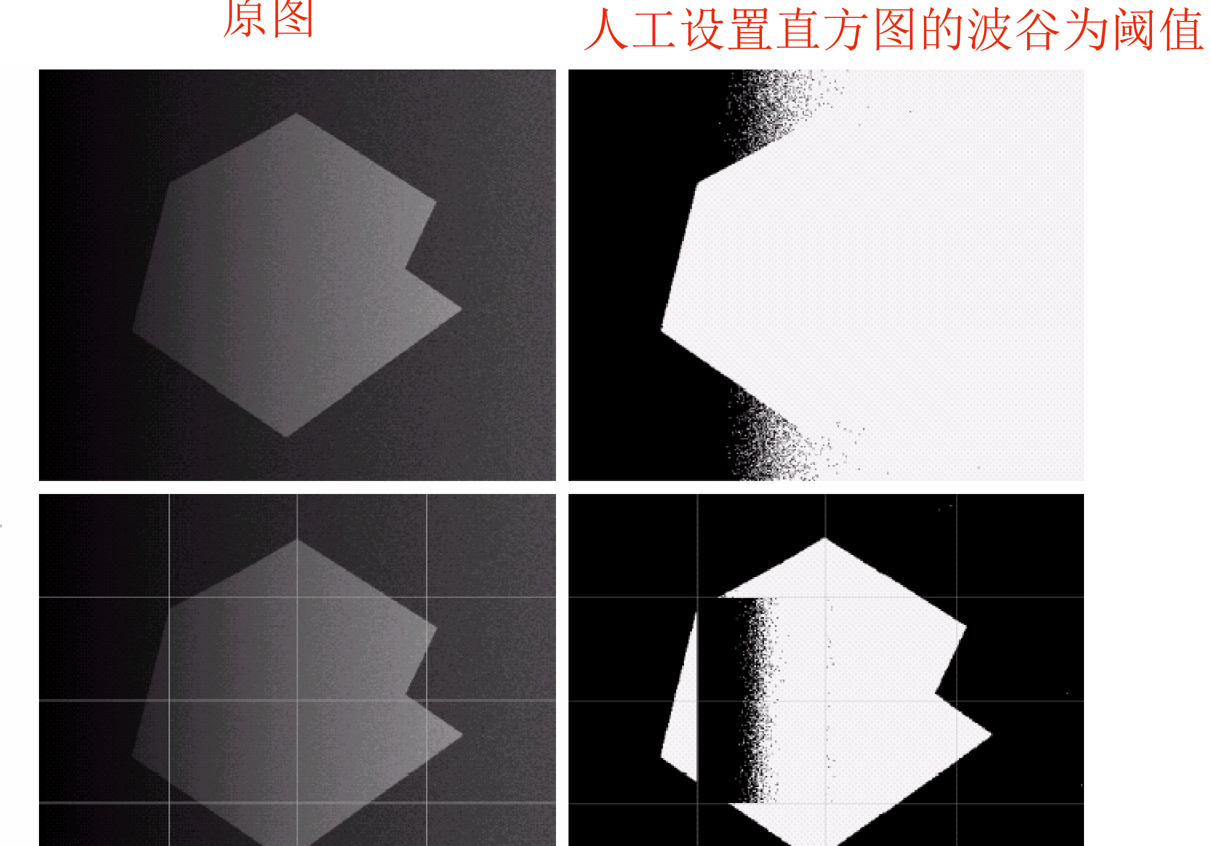


FIGURE 10.30 (a) Original image. (b) Result of global thresholding. (c) Image subdivided into individual subimages. (d) Result of adaptive thresholding.

分割为子图像：4等分后再4等分

自适应阈值处理的结果

图像分割



- 最佳全局和自适应阈值

  ✓ 假设一幅图像仅包含两个主要的灰度级区域。令z表示灰度级值，则两个灰度区域的直方图可以看作它们概率密度函数(PDF)的估计p(z)

  ✓ p(z)是两个密度的和或混合。一个是图像中亮区域的密度，另一个是暗区域的密度

  ✓ 如果p(z)已知或假设，则它能够确定一个最佳阈值（具有最低的误差）将图像分割为两个可区分的区域

最佳全局和自适应阈值



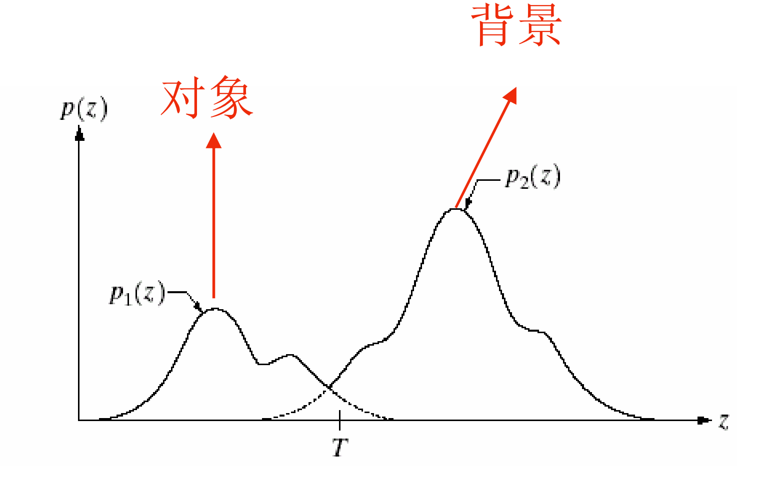
✓ 假设2个PDF中较大的一个对应背景的灰度级，较小的一个描述了图像中对象的灰度级，则混合PDF是

$$p(z) = P_1 p_1(z) + P_2 p_2(z)$$

P_1是属于对象像素的概率，P_2是属于背景像素的概率，假设图像只包括对象和背景，则

$$P_1 + P_2 = 1$$

FIGURE 10.32 Gray-level probability density functions of two regions in an image.



最佳全局和自适应阈值



- 在区间[a, b]内取值的随机变量的概率是它的概率密度函数从a到b的积分，即在这两个上下限之间PDF曲线围住的面积，因此，将一个背景点当作对象点进行分类时，错误发生的概率为：

$$E_1(T)=\int_{-\infty}^{T} p_2(z)dz$$

这是在曲线$p_2(z)$下方位于阈值左边区域的面积

- 将一个对象点当作背景点进行分类错误发生的概率为

$$E_2(T)=\int_{T}^{\infty} p_1(z)dz$$

这是在曲线$p_1(z)$下方位于阈值右边区域的面积

最佳全局和自适应阈值



- 出错率的整体概率是

$$E(T)=P_2E_1(T)+P_1E_2(T)$$

- 为了找到出错最少的阈值，使用莱布尼兹法则把E(T)对T求微分并令结果等于0，得到

$$P_2p_1(T)=P_1p_2(T)$$

- 上式解出T，即为最佳阈值
- 如果$P_1=P_2$，则最佳阈值位于曲线$p_1(z)$和$p_2(z)$的交点处

最佳全局和自适应阈值



- 高斯密度可以用两个参数均值和方差描述

$$p(z)=\frac{P_1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(z-\mu_1)^2}{2\sigma_1^2}}+\frac{P_2}{\sqrt{2\pi}\sigma_2}e^{-\frac{(z-\mu_2)^2}{2\sigma_2^2}}$$

- 出错最少的阈值T的解

$$AT^2+BT+C=0$$

$$A=\sigma_1^2-\sigma_2^2$$

$$B=2\left(\mu_1\sigma_2^2-\mu_2\sigma_1^2\right)$$

$$C=\sigma_1^2\mu_2^2-\sigma_2^2\mu_1^2+2\sigma_1^2\sigma_2^2\ln\left(\sigma_2P_1/\sigma_1P_2\right)$$

最佳全局和自适应阈值



- 如果方差相等 $\sigma^2=\sigma_1^2=\sigma_2^2$，则得到单一的阈值

$$T=\frac{\mu_1+\mu_2}{2}+\frac{\sigma^2}{\mu_1-\mu_2}\ln\left(\frac{P_2}{P_1}\right)$$

- 如果$P_1=P_2$，最佳阈值是均值的平均数

图像分割



- 通过边界特性选择阈值
  - ✓ 基本思想：
    - 如果直方图的各个波峰很高、很窄、对称，且被很深的波谷分开时，有利于选择阈值
    - 为了改善直方图的波峰形状，我们只把区域边缘的像素绘入直方图，而不考虑区域中间的像素
    - 用微分算子，处理图像，使图像只剩下边界中心两边的值

图像分割



- 通过边界特性选择阈值
  - ✓ 这种方法有以下优点：

    1)在前景和背景所占区域面积差别很大时，不会造成一个灰度级的波峰过高，而另一个过低

    2)边缘上的点在区域内还是区域外的概率是相等的，因此可以增加波峰的对称性

    3)基于梯度和拉普拉斯算子选择的像素，可以增加波峰的高度

图像分割



- 通过边界特性选择阈值
  - ✓ 算法的实现：

    1）对图像进行梯度计算，得到梯度图像。

    2）得到梯度值最大的那一部分（比如10%）的像素直方图

    3）通过直方图的谷底，得到阈值T
  - ✓ 如果用拉普拉斯算子，不通过直方图，直接得到阈值，方法是使用拉普拉斯算子过滤图像，将0跨越点对应的灰度值为阈值T

图像分割



- 基于不同变量的阈值

✓ 在某些情况下，传感器可以产生不止一个在图像中描述每一个像素的可利用的变量，因此，允许进行多谱段阈值处理

✓ 例如一幅有3个变量的图像(RGB分量)，每个像素有16种可能的灰度级，构成16×16×16种灰度级（网格，立方体）

✓ 阈值处理就是在三维空间内寻找点的聚簇的过程。如在直方图中找到有效点簇K，可以对RGB分量值接近某一个簇的像素赋予一个任意值(如白色的值)，对其它像素赋予另一个值（如黑色的值）

✓ 彩色图像处理中的色调和饱和度易于图像分割

基于不同变量阈值的处理

彩色照片的单色图像

对应于脸部色调的一个簇进行阈值处理得到

对应于红色轴的一个簇进行阈值处理得到，红色的围巾和花出现在分割结果中







a b c

FIGURE 10.39 (a) Original color image shown as a monochrome picture. (b) Segmentation of pixels with colors close to facial tones. (c) Segmentation of red components.

- ✓ 原彩色图是由16位RGB图像组成的
- ✓ 围巾是鲜红色，头发和脸部的颜色很浅

图像分割



- 概述
- 间断检测
- 边缘连接和边界检测
- 阈值处理
- 基于区域的分割
- 分割中运动的应用

图像分割



- 基于区域的分割
  - ✓ 基本公式
  - ✓ 区域生长
  - ✓ 区域分离与合并

图像分割



- 基本概念
  - ✓ 目标：将区域R划分为若干个子区域$R_1, R_2, \ldots, R_n$，这些子区域满足5个条件：

    1) 完备性：$\bigcup_{i=1}^{n} R_i = R$

    2) 连通性：每个R_i都是一个连通区域

    3) 独立性：对于任意$i \neq j$，$R_i \cap R_j = \Phi$

图像分割

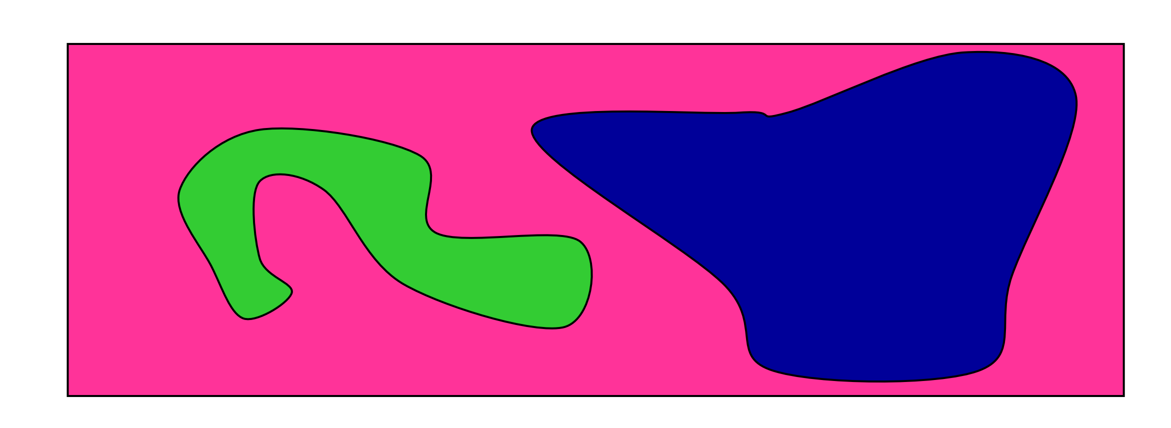


- 基本概念

  4) 单一性：每个区域内的灰度级相等，

  $P(R_i) = TRUE,\ i = 1, 2, \ldots, n$

  5) 互斥性：任两个区域的灰度级不等，$P(R_i \cup R_j) = FALSE,\ i \neq j$



图像分割



- 区域增长的算法实现：

  1）根据图像的不同应用选择一个或一组种子，它或者是最亮或最暗的点，或者是位于点簇中心的点

  2）选择一个描述符（条件）

  3）从该种子开始向外扩张，首先把种子像素加入结果集合，然后不断将与集合中各个像素连通、且满足描述符的像素加入集合

  4）上一过程进行到不再有满足条件的新结点加入集合为止



图像分割

- 通过像素集合的区域增长
  - ✓ 算法实现：

区域A

种子像素

区域B

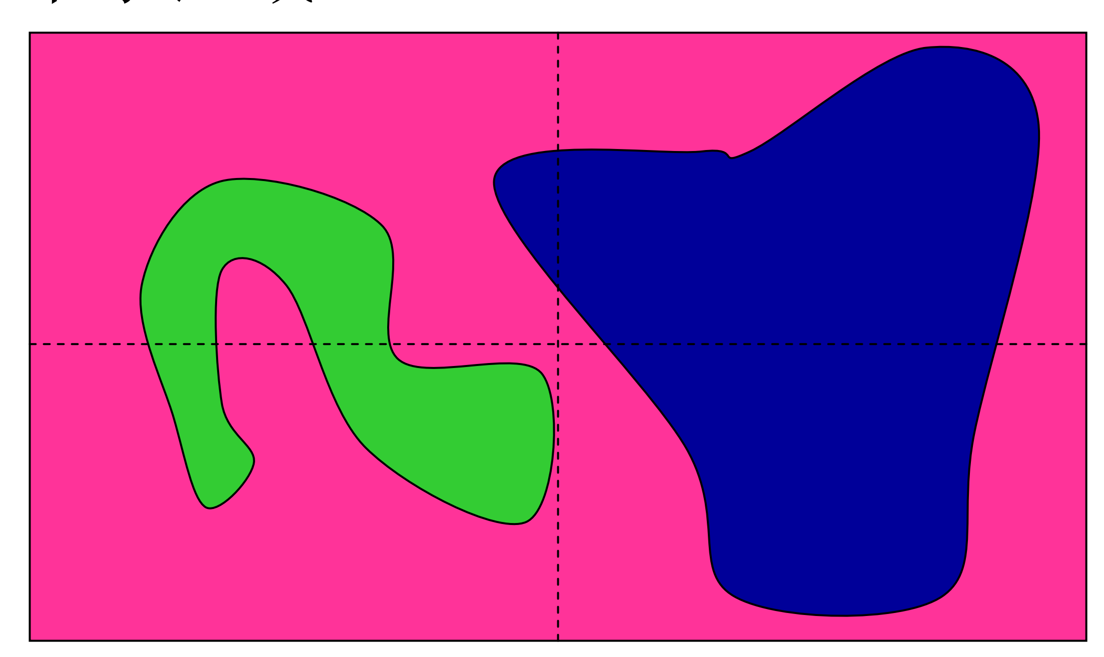
种子像素

图像分割



- 区域分裂与合并
  - 算法实现：

    1）对图像中灰度级不同的区域，均分为四个子区域



图像分割



- 区域分裂与合并
  - ✓算法实现：

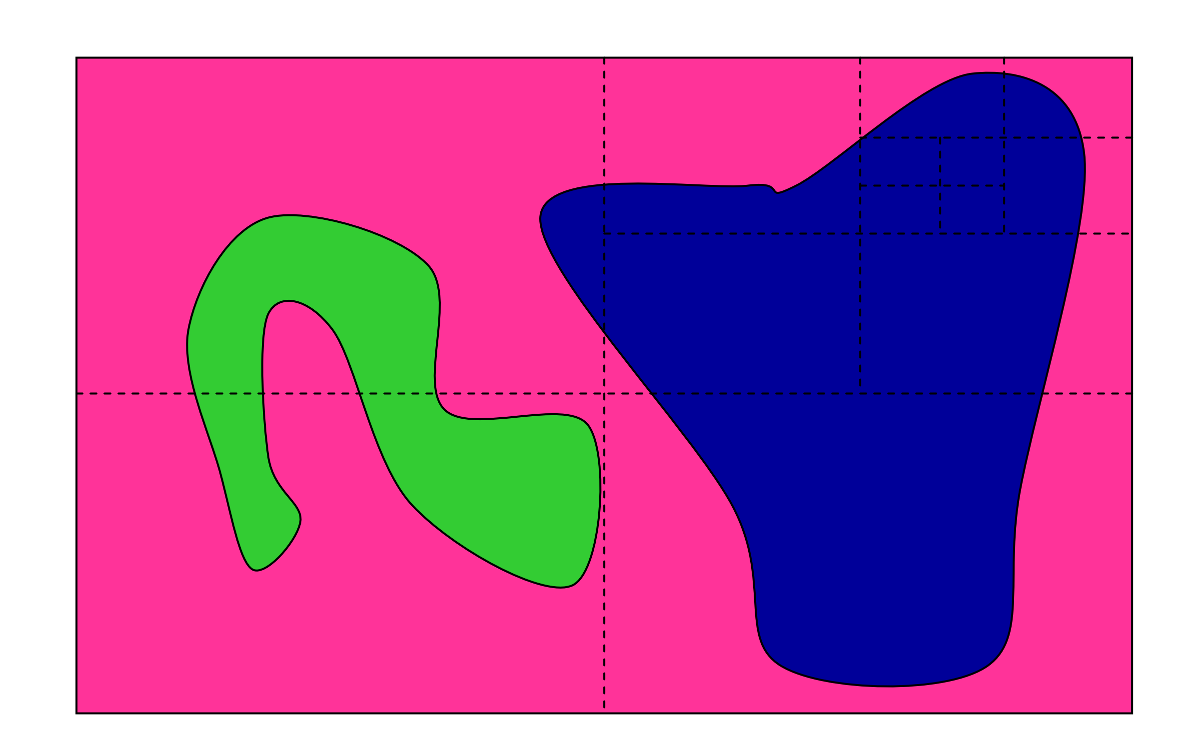
    2）如果相邻的子区域所有像素的灰度级相同，则将其合并

    3）反复进行上两步操作，直至不再有新的分裂与合并为止

图像分割



- 区域分裂与合并
  - ✓ 算法实现：



图像分割



- 区域分裂与合并
  - ✓ 算法实现：实际应用中还可作以下修改：

    $P(R_i)$的定义为：

    1）区域内多于80%的像素满足不等式

    $$|z_j - m_i| <= 2\sigma_i,$$

    其中：z_j是区域R_i中第j个点的灰度级，

    m_i是该区域的平均灰度级，

    σ_i是区域的灰度级的标准方差。

    2）当$P(R_i)$=TRUE时，将区域内所有像素的灰度级置为m_i

图像分割



- 概述
- 间断检测
- 边缘连接和边界检测
- 阈值处理
- 基于区域的分割
- 分割中运动的应用

分割中运动的应用



- 空间技术

✓ 使用两帧图像$f(x,y,t_i)$和$f(x,y,t_j)$相减的办法，形成差值图像

$$d_{ij}(x,y)=\begin{cases}1 & \left|f(x,y,t_i)-f(x,y,t_j)\right|>T \\ 0 & \text{其它}\end{cases}$$

✓ 在动态图像处理过程中，d_{ij}中值为1的像素被认为是对象运动的结果

✓ 考虑图像帧序列f(x,y,t1),f(x,y,t2),...,f(x,y,tn),并令f(x,y,t1)为基本图像，一幅累积差异图像(ADI)由基准图像和图像序列的后续图像对比得到

分割中运动的应用



- 空间技术（续）

  ✓ 令R(x, y)表示基准图像，绝对ADI，正ADI和负ADI定义如下：

$$A_k(x,y)=\begin{cases}A_{k-1}(\mathrm{x,y})+1 & |R(x,y)-f(x,y,k)|>T \\ A_{k-1}(\mathrm{x,y}) & \text{其它}\end{cases}$$

$$P_k(x,y)=\begin{cases}P_{k-1}(\mathrm{x,y})+1 & [R(x,y)-f(x,y,k)]>T \\ P_{k-1}(\mathrm{x,y}) & \text{其它}\end{cases}$$

$$N_k(x,y)=\begin{cases}N_{k-1}(\mathrm{x,y})+1 & [R(x,y)-f(x,y,k)]<-T \\ N_{k-1}(\mathrm{x,y}) & \text{其它}\end{cases}$$

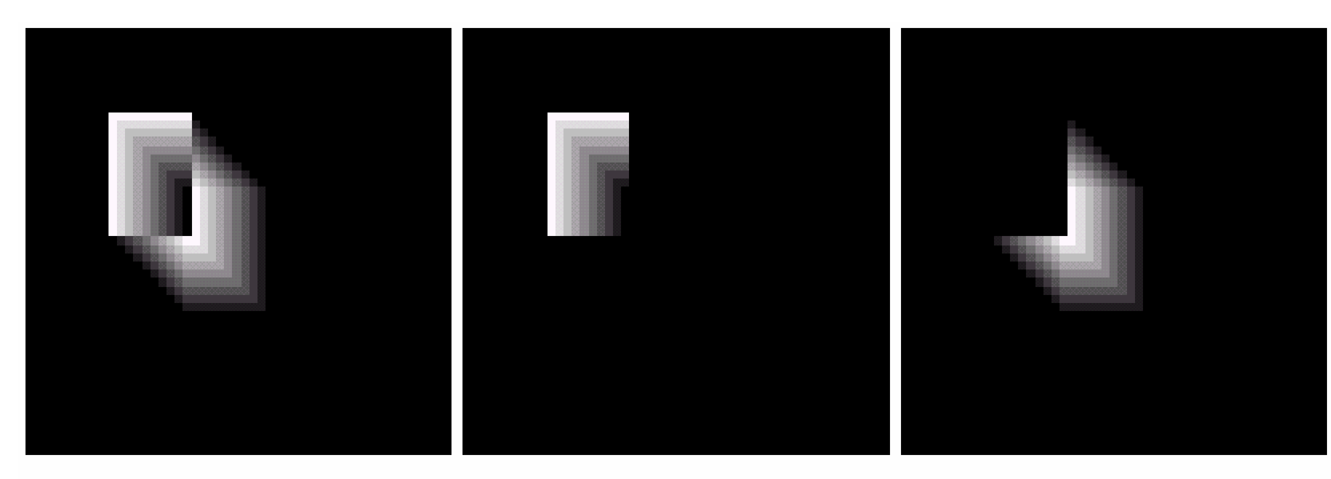
分割中运动的应用



- 空间技术举例

  向东南方向运动的矩形目标的ADI

绝对ADI　　正ADI　　负ADI



a b c

FIGURE 10.49 ADIs of a rectangular object moving in a southeasterly direction. (a) Absolute ADI. (b) Positive ADI. (c) Negative ADI.

数字图像处理（11）



任何问题？

研究生课程



数字图像处理

Digital Image Processing

彭宇新

北京大学计算机科学技术研究所

E_mail:pengyuxin@icst.pku.edu.cn

表示与描述



- 概述
- 表示方法
- 边界描述子
- 关系描述子

表示与描述



- 概述

  ✓ 图像分割结果是得到了区域内的像素集合，或位于区域边界上的像素集合，这两个集合是互补的

  ✓ 与分割类似，图像中的区域可用其内部(如组成区域的像素集合)表示，也可用其外部（如组成区域边界的像素集合）表示

  ✓ 一般来说，如果关心的是区域的反射性质，如灰度、颜色、纹理等，常用内部表示法；如果关心的是区域形状，则选用外部表示法

  ✓ 表示是直接具体地表示目标，好的表示方法应具有节省存储空间、易于特征计算等优点

表示与描述



- 概述

  ✓ 描述是较抽象地表示目标。好的描述应在尽可能区别不同目标的基础上对目标的尺度、平移、旋转等不敏感，这样的描述比较通用

  ✓ 描述可分为对边界的描述和对区域的描述。此外，边界和边界或区域和区域之间的关系也常需要进行描述

  ✓ 表示和描述是密切联系的。表示的方法对描述很重要，因为它限定了描述的精确性；而通过对目标的描述，各种表示方法才有实际意义

  ✓ 表示和描述又有区别，表示侧重于数据结构，而描述侧重于区域特性以及不同区域间的联系和差别

表示与描述



- 概述

  ✓ 对目标特征的测量是要利用分割结果进一步从图像中获取有用信息，为达到这个目的需要解决两个关键问题：

    - 选用什么特征来描述目标
    - 如何精确地测量这些特征

  ✓ 常见的目标特征分为灰度（颜色）、纹理和几何形状特征等。其中，灰度和纹理属于内部特征，几何形状属于外部特征

表示与描述



- 概述
- 表示方法
- 边界描述子
- 关系描述子

表示与描述



- 表示方法
  - ✓ 链码
  - ✓ 多边形近似
  - ✓ 外形特征
  - ✓ 边界分段
  - ✓ 区域骨架

表示与描述



- 链码
  - ✓ 链码用于表示由顺序连接的具有指定长度和方向的直线段组成的边界线
  - ✓ 这种表示方法基于线段的4或8连接
  - ✓ 每一段的方向使用数字编号方法进行编码

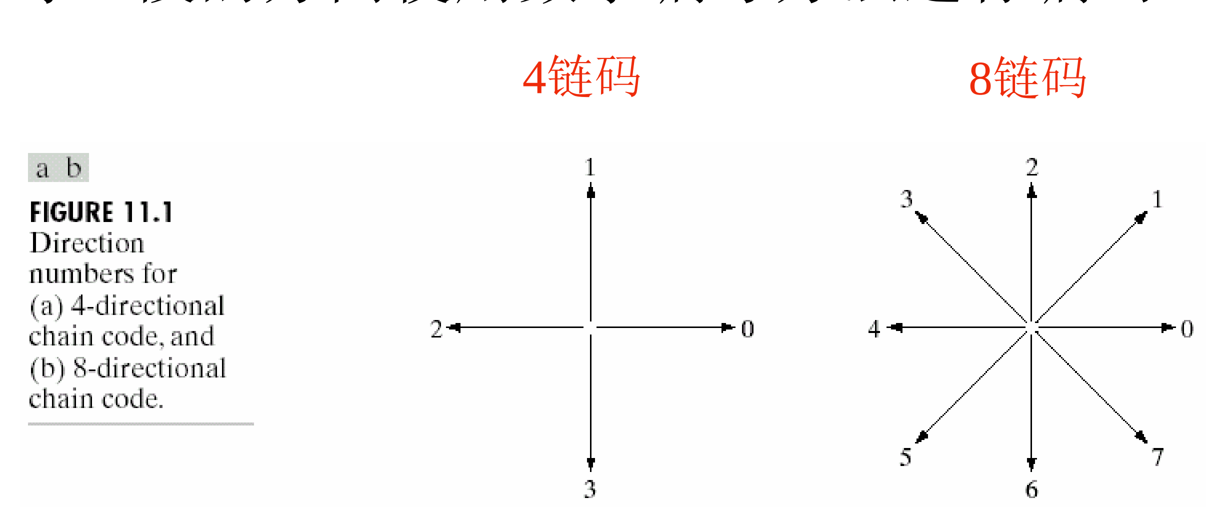
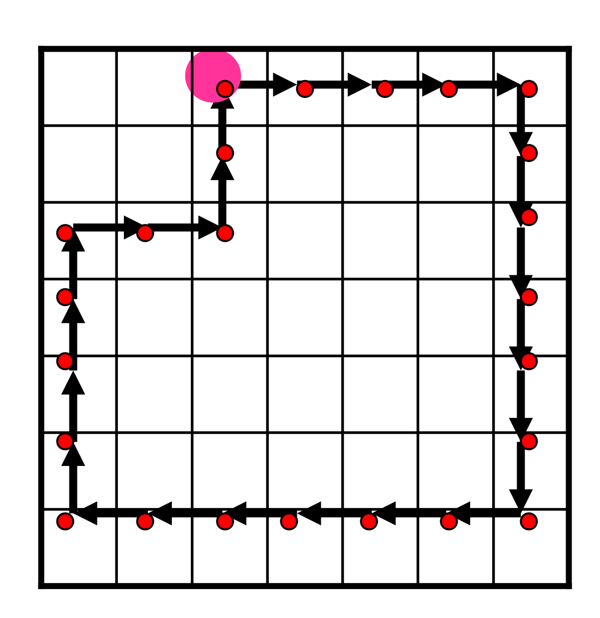


FIGURE 11.1 Direction numbers for (a) 4-directional chain code, and (b) 8-directional chain code.

表示与描述



- 链码举例：



4-链码：00003333332222221111110011

表示与描述



- 链码
  - ✓ 算法：
    - 给每一个线段边界一个方向编码
    - 有4链码和8链码两种编码方法
    - 从起点开始，沿边界编码，至起点被重新碰到，结束一个对象的编码

表示与描述



- 链码
  - ✓ 问题1：

    1）链码相当长

    2）噪音会产生不必要的链码
  - ✓ 改进1：

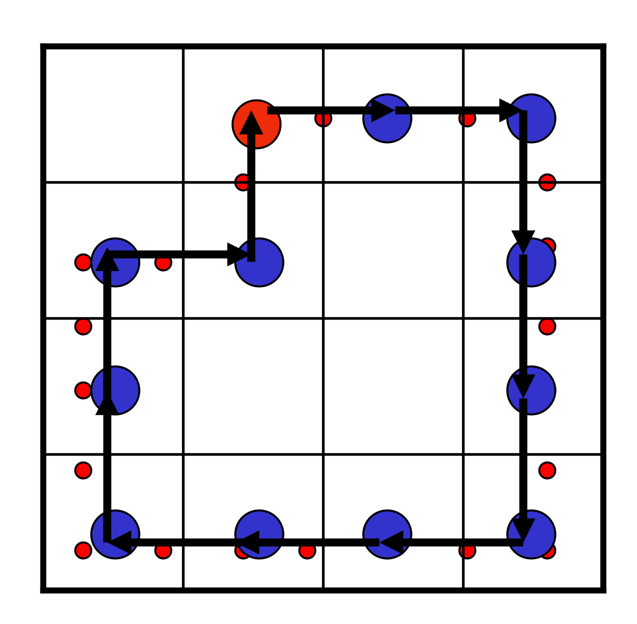
    1）加大网格空间

    2）依据原始边界与结果的接近程度，来确定新点的位置

表示与描述



- 链码举例：



4-链码：003332221101

表示与描述



- 链码
  - ✓ 问题2：
    - 1）由于起点的不同，造成编码的不同
    - 2）由于角度的不同，造成编码的不同
  - ✓ 改进2：
    - 1）从固定位置作为起点(最左最上)开始编码
    - 2）通过使用链码的差分代替码字本身的方式

表示与描述



- 链码
  - ✓ 循环差分链码：用相邻链码的差代替链码

例如：4-链码　10103322

循环差分为：　33133030

循环差分：
$1 - 2 = -1(3)$　　$3 - 0 = 3$
$0 - 1 = -1(3)$　　$3 - 3 = 0$
$1 - 0 = 1$　　$2 - 3 = -1(3)$
$0 - 1 = -1(3)$　　$2 - 2 = 0$

表示与描述



- 多边形近似
  - ✓ 基本思想：用最少的多边形线段，获取边界形状的本质
  - ✓ 寻找最小基本多边形的方法一般有两种：

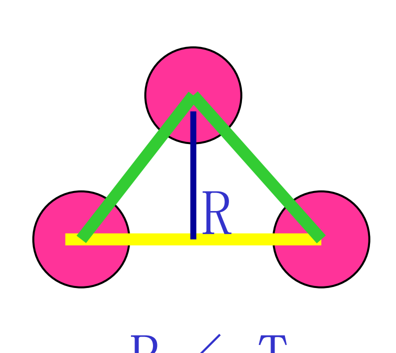
    1）点合成法

    2）边分裂法

表示与描述



- 多边形近似



✓ 点合成算法：

1）沿着边界选两个相邻的点对，计算首尾连接直线段与原始折线段的误差R。

2）如果误差R小于预先设置的阈值T。去掉中间点，选新点对与下一相邻点对，重复1）；否则，存储线段的参数，置误差为0，选被存储线段的终点为起点，重复1）2）。

3）当程序的第一个起点被遇到，算法结束。

表示与描述



- 多边形近似
  - ✓ 点合成算法思想举例：

R

R < T

表示与描述



- 多边形近似
  - ✓ 点合成算法的问题：
    - ➢ 顶点一般不对应于边界的拐点（如拐角）。因为新的线段直到超过误差的阈值才开始画

      例如：如果沿着一条长的直线追踪，而它出现了一个拐角，在超过阈值之前，拐角上的一些点会被丢弃
    - ➢ 下面讲到的分裂法可用于缓解这个问题

表示与描述



- 多边形近似
  - ✓ 分裂边算法：

    （1）连接边界线段的两个端点（如果是封闭边界，连接最远点）；

    （2）如果最大正交距离大于阈值，将边界分为两段，最大值点定位一个顶点。重复（1）；

    （3）如果没有超过阈值的正交距离，结束。

表示与描述



- 多边形近似
  - ✓ 边分裂算法思想举例：

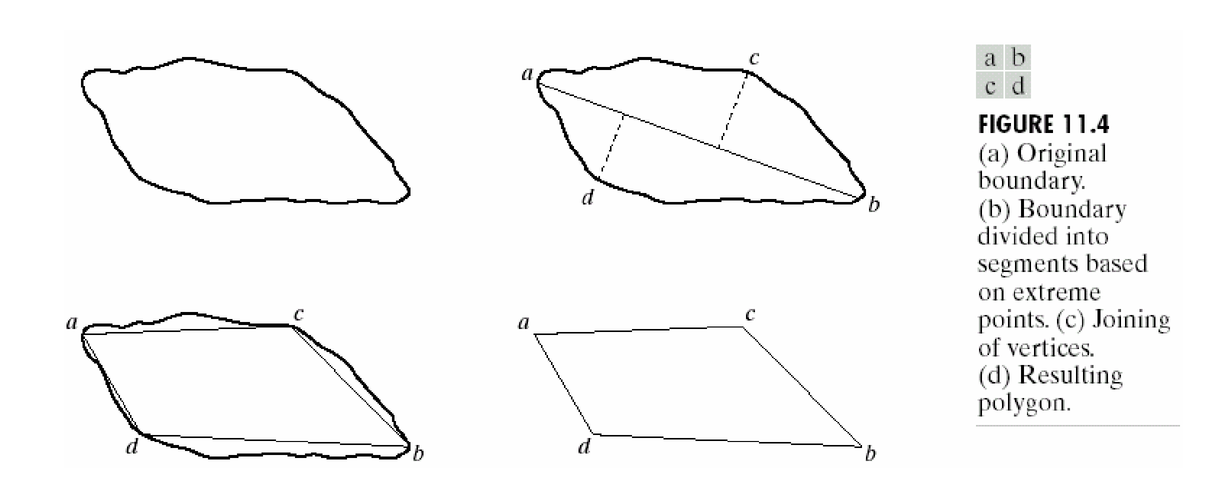


FIGURE 11.4 (a) Original boundary. (b) Boundary divided into segments based on extreme points. (c) Joining of vertices. (d) Resulting polygon.

使用直线ab长度的0.25倍作为阈值的拆分过程结果。由于在新的边界线段上没有超过阈值的垂直距离的点，分割过程终止

表示与描述



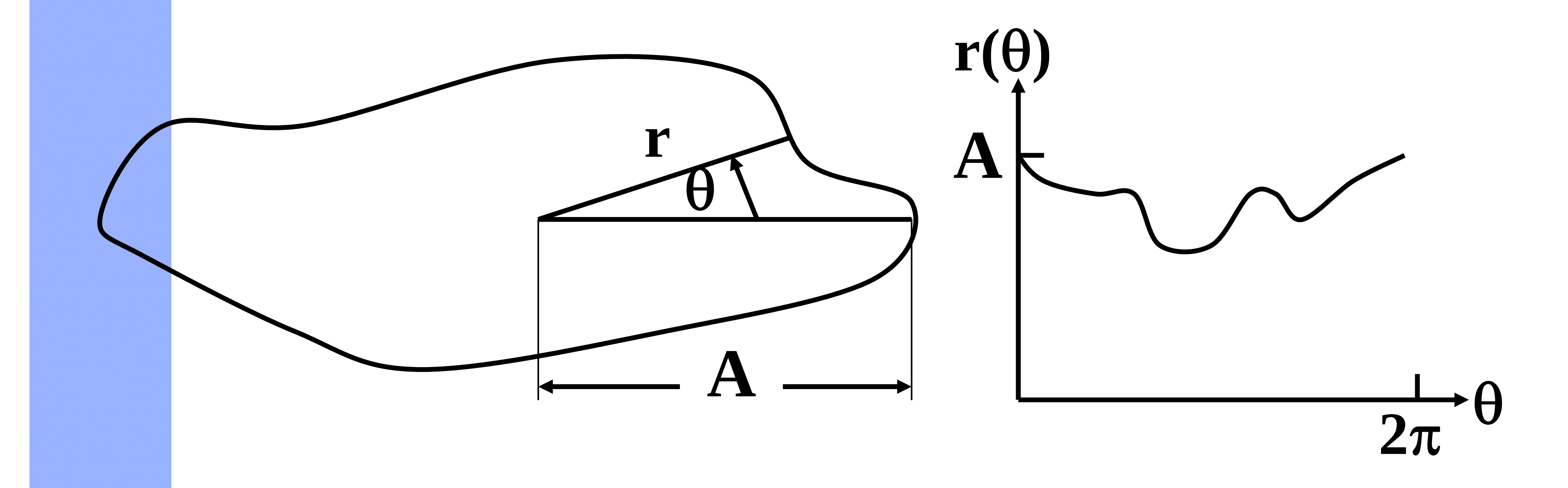
- 外形特征
  - ✓ 基本思想：

    外形特征是一种用一维函数表达边界的方法。基本思想是把边界的表示降到一维函数

表示与描述



- 外形特征
  - ✓ 函数定义——质心角函数：边上的点到质心的距离r，作为夹角θ的函数r(θ)



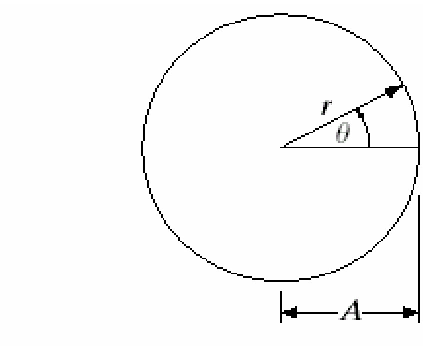
外形特征举例

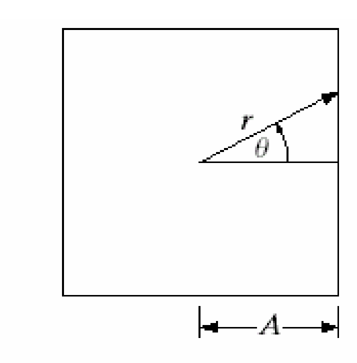


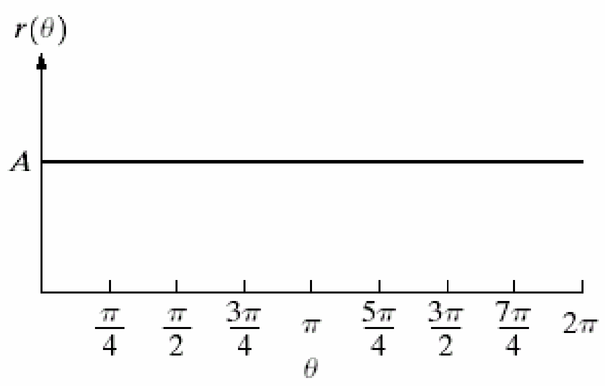
到达正方形的4个对角上达到最大值

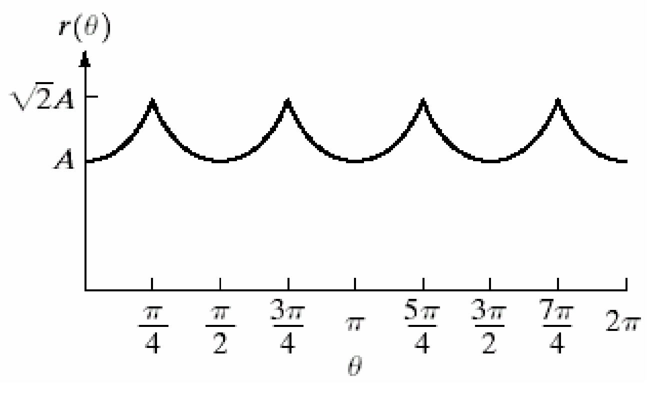
a b

FIGURE 11.5 Distance-versus-angle signatures. In (a) $r(\theta)$ is constant. In (b), the signature consists of repetitions of the pattern $r(\theta) = A\sec\theta$ for $0 \le \theta \le \pi/4$ and $r(\theta) = A\csc\theta$ for $\pi/4 < \theta \le \pi/2$.









表示与描述



- 外形特征
  - 问题：函数依赖于旋转和比例缩放变换
  - 改进：
    - 对于旋转——两种改进：

      a. 选择离质心最远的点作为起点

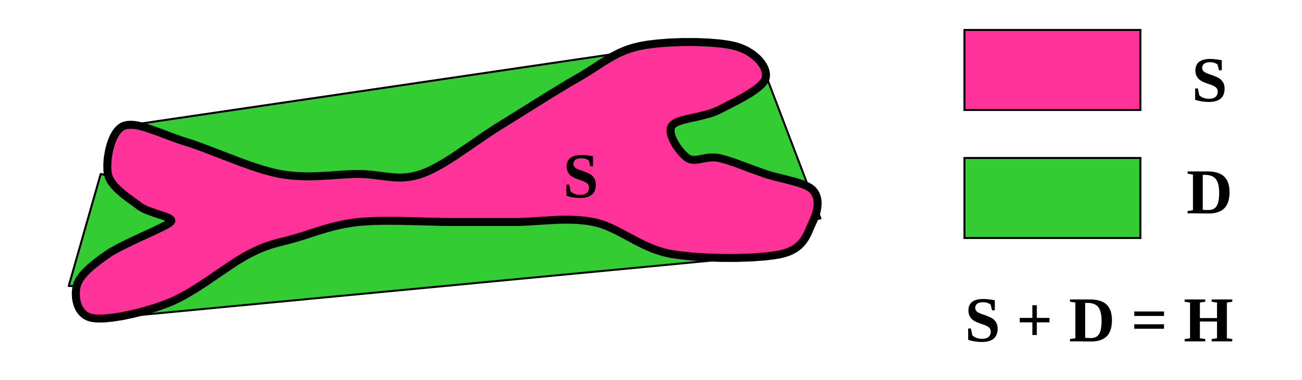
      b. 选择从质心到主轴最远的点作为起点
    - 对于比例变换：

      对函数进行正则化，使函数值总是分布在相同的值域里，比如说[0，1]

表示与描述



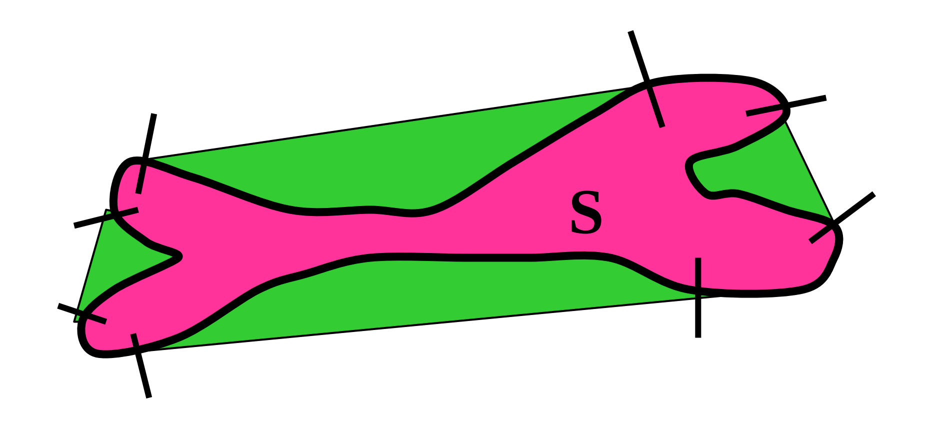
- 边界分段
  - ✓ 基本概念：
    - ➢ 一个任意集合S（区域）的凸起外缘H是：包含S的最小凸起的集合
    - ➢ H-S的差的集合被称为集合S的凸起补集D



表示与描述



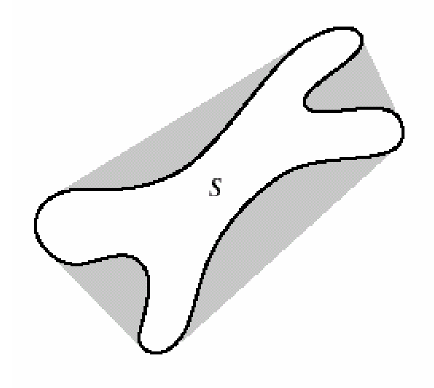
- 边界分段
  - ✓ 分段算法：
    - 给进入和离开凸起补集D的变换点打标记来划分边界段。
    - 优点：不依赖于方向和比例的变化

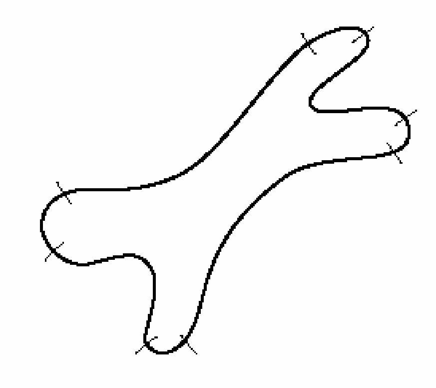


表示与描述



- 边界分段举例





a b

FIGURE 11.6
(a) A region, *S*, and its convex deficiency (shaded).
(b) Partitioned boundary.

区域S和它的凸起补集D　　　被分割的边界

表示与描述



- 边界分段
  - ✓ 问题：

    噪音的影响，导致出现零碎的划分。
  - ✓ 解决的方法：

    先平滑边界，或用多边形逼近边界，然后再分段

表示与描述



- 区域骨架
  - ✓ 基本思想
    - 表示一个平面区域结构形状的重要方法是把它削减成图形。这种削减可以通过细化（也称为抽骨架）算法，获取区域的骨架来实现
    - Blum的中轴变换方法（MAT）

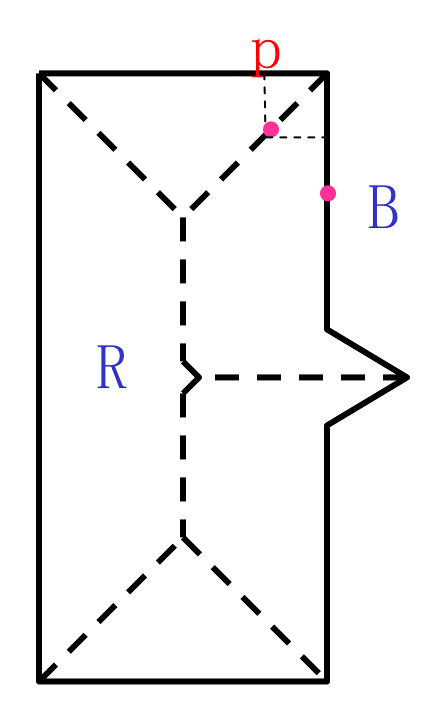
      设:R是一个区域，B为R的边界点，对于R中的点p，找p在B上“最近”的邻居。如果p有多于一个的邻居，称它属于R的中轴（骨架）

表示与描述



- 区域骨架（续）
  - ✓ 问题：计算量大

    包括计算区域的每个内部点到其边界点的距离



表示与描述



- 区域骨架
  - ✓ 算法改进思想

    在保证产生正确骨架的同时，改进算法的效率。比较典型的是一类细化算法，它们不断删除区域边界点，但保证删除满足：

    （1）不删除端点

    （2）不破坏连通性

    （3）不造成对区域的过度腐蚀

表示与描述



- 区域骨架
  - ✓ 一种细化二值区域的算法
    - ➢ 假设区域内的点值为1，背景值为0
    - ➢ 这个方法由对给定区域的边界点连续进行两个基本操作构成
    - ➢ 这里边界点是指任何值为1且至少有一个8邻域上的点为0的像素

表示与描述



- 区域骨架

p9	p2	p3
p8	p1	p4
p7	p6	p5

p9	p2	p3
p8	p1	p4
p7	p6	p5

p9	p2	p3
p8	p1	p4
p7	p6	p5

  - ✓ 基本操作1

对于满足以下四个条件的边界点打标记准备删除：

(a) $2 \le N(p_1) \le 6$ 其中$N(p_1)$是点p_1的邻域中1的个数，

即： $N(p_1)=p_2+p_3+\ldots+p_9$

(b) $S(p_1) = 1$

其中$S(p_1)$是按$p_2, p_3, \ldots, p_9$顺序，0-1转换的个数

(c) $p_2 . p_4 . p_6 = 0$ （p_2 , p_4 , p_6 至少有一个0）

(d) $p_4 . p_6 . p_8 = 0$ （p_4 , p_6 , p_8 至少有一个0）

表示与描述



- 区域骨架

  所有条件都满足，才打删除标记。删除并不立即进行，而是等到对所有边界点都打完标记后，再把作了标记的点一起删除

  ✓ 举例：

0	0	1
1	p1	0
1	0	1

p9	p2	p3
p8	p1	p4
p7	p6	p5

p9	p2	p3
p8	p1	p4
p7	p6	p5

$N(p_1) = 4$

$S(p_1) = 3$

$p_2 . p_4 . p_6 = 0$

$p_4 . p_6 . p_8 = 0$　第2个条件没满足不打标记

表示与描述



p9	p2	p3
p8	p1	p4
p7	p6	p5

p9	p2	p3
p8	p1	p4
p7	p6	p5

- 区域骨架
  - ✓ 基本操作2

    条件(a)、(b)与操作1相同

    条件(c)、(d)改为：

    c’) $p_2 \cdot p_4 \cdot p_8 = 0$

    d’) $p_2 \cdot p_6 \cdot p_8 = 0$

表示与描述



- 区域骨架
  - ✓ 细化算法

    细化算法的一轮操作包括：

    - 按操作1，给边界点打标记——删除点
    - 按操作2，给边界点打标记——删除点
    - 这个基本过程反复进行，直至没有点可以删除为止。此时算法终止。

表示与描述



- 概述
- 表示方法
- 边界描述子
- 关系描述子

表示与描述



- 边界描述子
  - ✓ 简单描述子
  - ✓ 形状数
  - ✓ 傅里叶描述子
  - ✓ 矩量

表示与描述



- 简单描述子
  - ✓ 边界的周长：

    是最简单的描述符之一。沿轮廓线计算像素的个数，给出了一个长度的近似估计
  - ✓ 边界的直径：边界B的直径是：

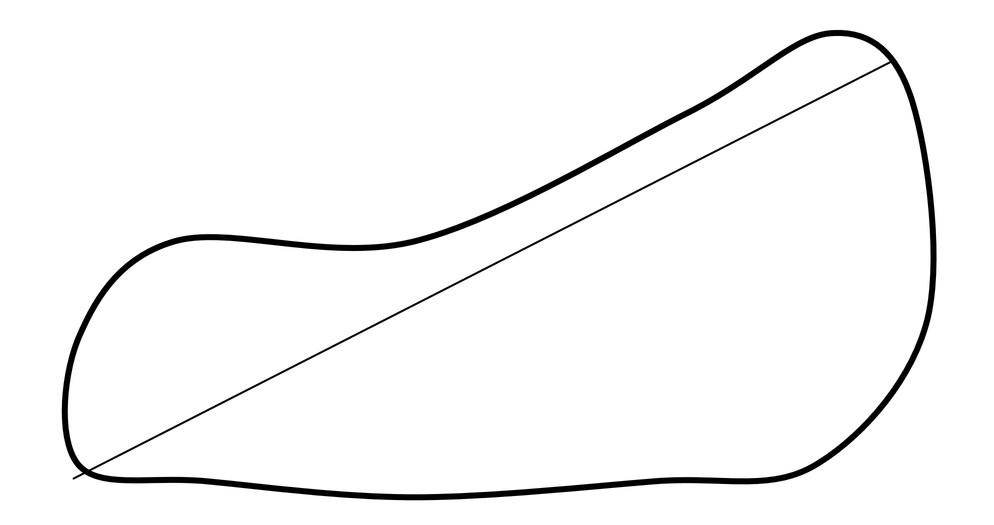
$$Diam(B) = \max_{i,j}\left[D\left(p_i, p_j\right)\right]$$

D是欧氏距离或几何距离，p_i，p_j是边界上的点。直径的长度和直径的两个端点连线（这条线被称为边界的主轴）的方向，是关于边界的有用的描述符。

表示与描述

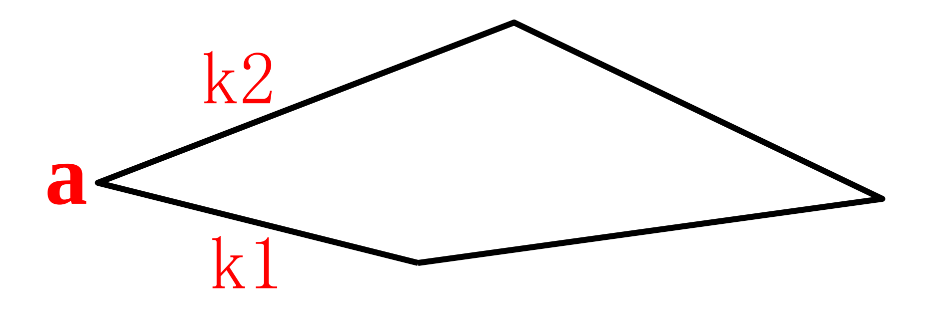


- 简单描述子
  - ✓ 边界的直径举例



表示与描述





- 简单描述子
  - ✓ 边界的曲率：

曲率被描述为斜率的变化率。近似：

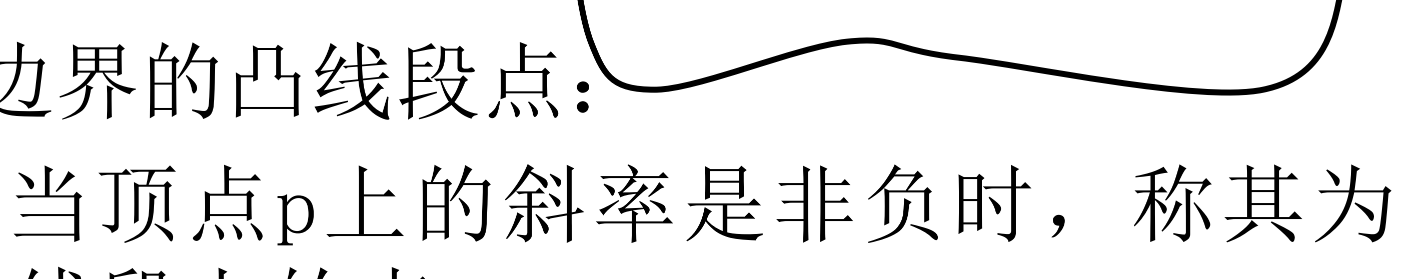
用相邻边界线段（描述为直线）的斜率差作为在边界线交点处的曲率描述子。

交点a处的曲率为　　$dk = k1 - k2$

其中k1、k2 为相邻线段的斜率

表示与描述





- 简单描述子
  - ✓ 边界的凸线段点：

    当顶点p上的斜率是非负时，称其为凸线段上的点
  - ✓ 边界的凹线段点：

    当顶点p上的斜率为负时，称其为凹线段上的点

表示与描述



- 形状数——链码的实用化
  - ✓ 形状数定义：最小循环首差链码

    循环首差链码：用相邻链码的差代替链码

    例如：4链码 10103322

    循环首差为：33133030

    循环首差:1 - 2 = -1(3)　　3 - 0 = 3

    0 - 1 = -1(3)　　3 - 3 = 0

    1 - 0 = 1　　2 - 3 = -1(3)

    0 - 1 = -1(3)　　2 - 2 = 0



表示与描述



- 形状数
  - ✓ 形状数定义：最小循环首差链码

例如：　4-链码　：10103322

循环首差　：33133|030

形状数　：03033133

  - ✓ 形状数序号n的定义：

形状数表达形式中的位数。上例序数为8

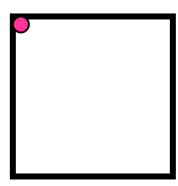
对于封闭边界序号一定是偶数。如4、6、8。



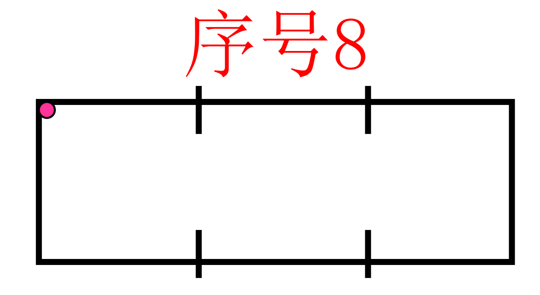
表示与描述

- 形状数
  - ✓ 序号为4、6、8的形状数举例：





序号6



链码：0321
首差：3333
形状：3333

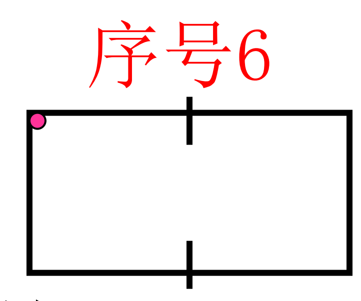
链码：003221
首差：303303
形状：033033

链码：00032221
首差：30033003
形状：00330033

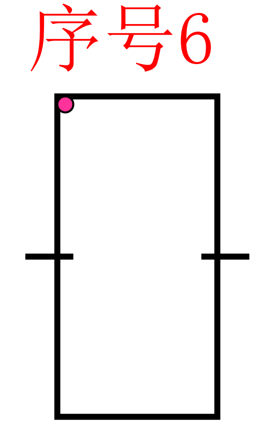
表示与描述



- 形状数
  - ✓ 序号为6的形状数举例：



链码： 003221
首差： 303303
形状： 033033



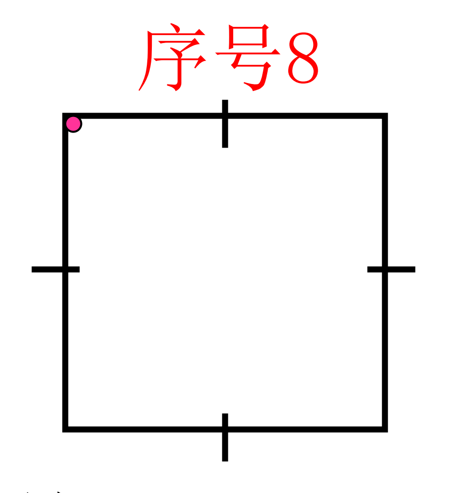
链码： 033211
首差： 330330
形状： 033033

形状数与方向无关

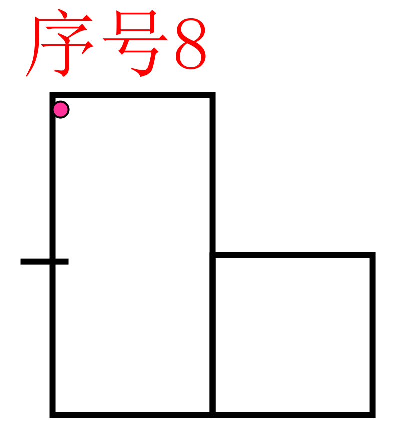
表示与描述



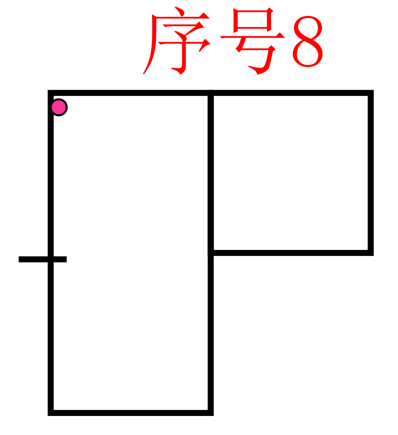
- 形状数
  - ✓ 序号为8的形状数举例：



链码：00332211
首差：30303030
形状：03030303



链码：03032211
首差：33133030
形状：03033133



链码：00323211
首差：30331330
形状：03033133

表示与描述



- 形状数
  - ✓ 问题：

    虽然链码的首差是不依赖于旋转的，但一般情况下边界的编码依赖于网格的方向。
  - ✓ 改进：

    规整化网格方向，具体方法如下：

表示与描述



- 形状数
  - 几个基本概念：
    - 边界最大轴a:是连接距离最远的两个点的线段
    - 边界最小轴b:与最大轴垂直，且其长度确定的包围盒刚好包围边界。
    - 边界离心率c：最大轴长度与最小轴长度的比　　$c = a / b$
    - 基本矩形：包围边界的矩形。

表示与描述



- 形状数
  - ✓ 基本概念举例

表示与描述



- 形状数
  - ✓ 规整化网格方向算法的思想：

    大多数情况下，将链码网格与基本矩形对齐，即可得到一个唯一的形状数。

    规整化网格方向的一种算法如下 ：

    （1）首先确定形状数的序号n；

    （2）在序号为n的矩形形状数中，找出一个与给定形状的基本矩形的离心率最接近的形状数

表示与描述



- 形状数

（3）然后再用这个矩形与基本矩形对齐，构造网格。

（4）用获得链码的方法得到链码；

（5）再得到循环首差；

（6）首差中的最小循环数即为形状数。

例：如果n=12，所有序号为12的矩形（即周长为12）为2×4，3×3，1×5。如果2×4矩形的离心率最接近于给定边界的基本矩形的离心率，我们建立一个2×4的网格。

表示与描述



- 形状数
  - ✓ 规整化网格方向算法举例：

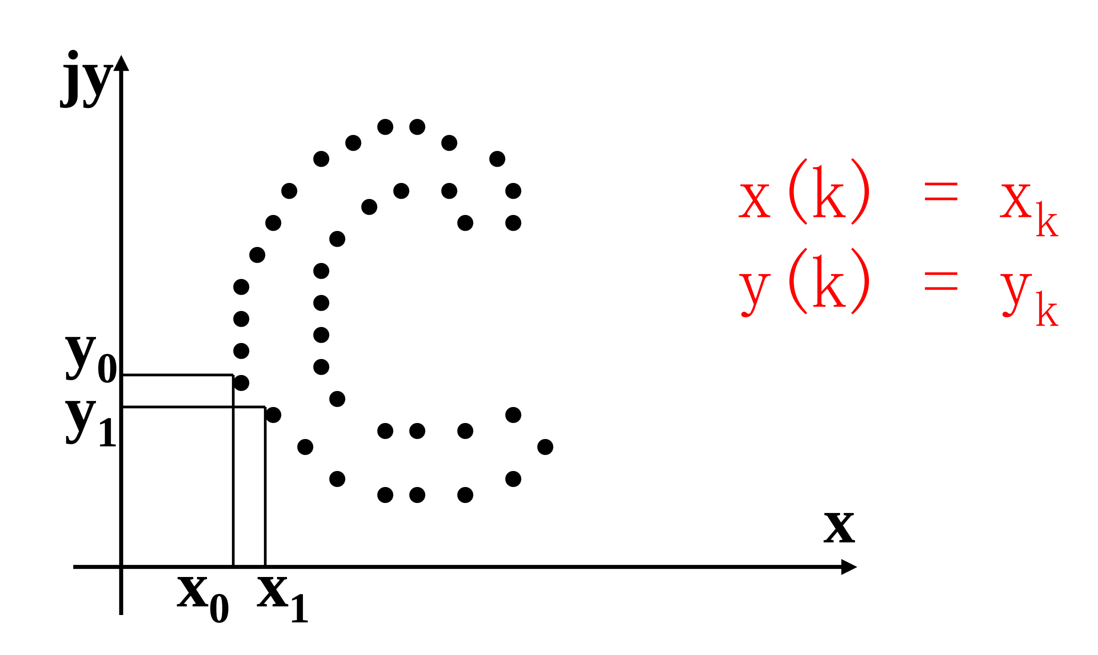
链码：000033222121
首差：300030300313
形状：003030031 33

表示与描述



- 傅里叶描述子：将一个二维问题简化成一个一维问题
  - ✓ 1）基本思想：

    （1）对于XY平面上的每个边界点，将其坐标用复数表示为：　$s(k) = x(k) + jy(k)$，　$k=0, 1, \ldots, N-1$



表示与描述



- 傅里叶描述子
  - 1）基本思想：

（2）进行离散傅里叶变换

$$a(u) = 1/N \sum_{u=0}^{N-1} s(k)\exp(-j2\pi uk/N) \qquad u=0, 1, \ldots, N-1$$

$$s(k) = \sum_{u=0}^{N-1} a(u)\exp(j2\pi uk/N) \qquad k=0, 1, \ldots, N-1$$

系数a(u)被称为边界的傅里叶描述子

表示与描述



- 傅里叶描述子
  - ✓ 1）基本思想：

（3）选取整数 $P \le N-1$，进行逆傅里叶变换（重构）

$$s'(k) = \sum_{u=0}^{P-1} a(u)\exp(j2\pi uk/N) \qquad k=0,1,\ldots,N-1$$

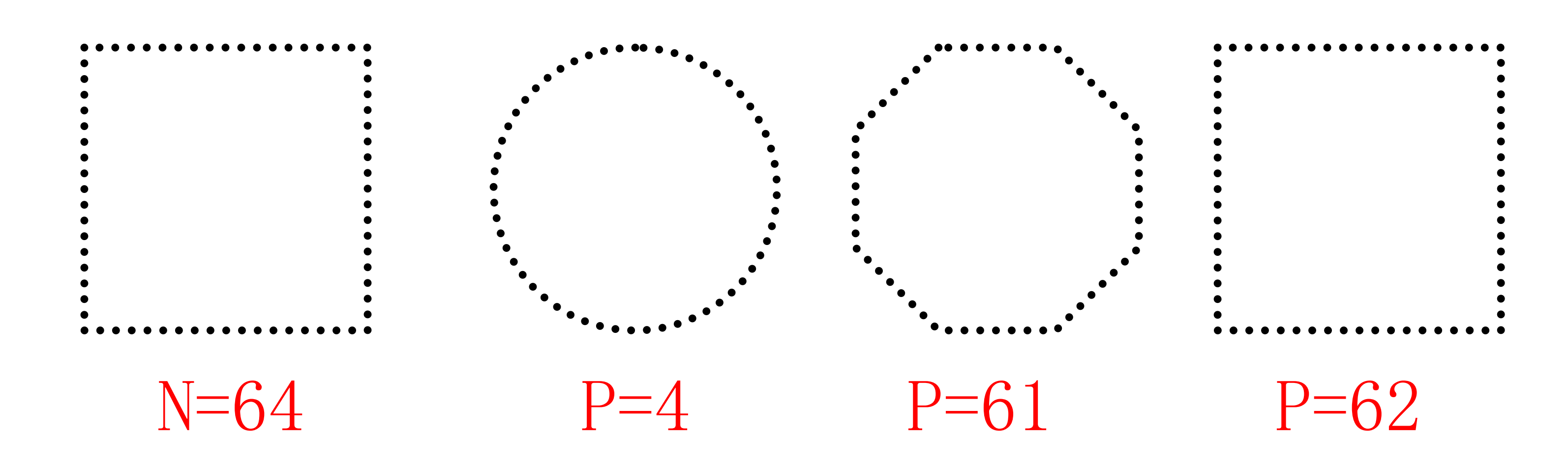
这时，对应于边界的点数没有改变，但在重构每一个点所需要的计算项大大减少了。如果边界点数很大，P一般选为2的指数次方的整数。

表示与描述



- 傅里叶描述符
  - 2）P的选取与描述符的关系

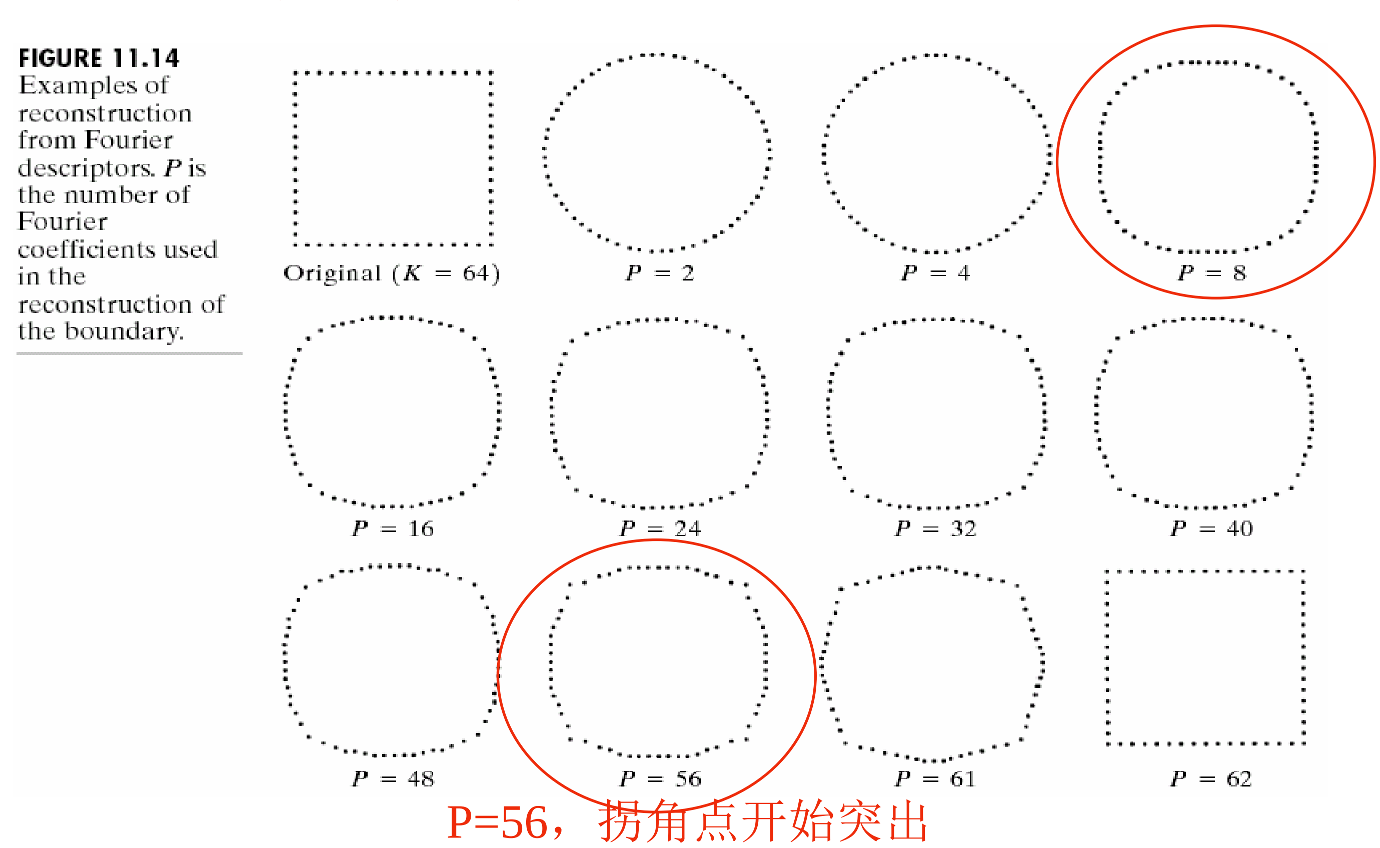
在上述方法中，相当于对于u ＞ P-1的部分舍去不予计算。由于傅里叶变换中高频部分对应于图像的细节描述，因此P取得越小，细节部分丢失得越多。



表示与描述



原图的方形边界



P=56，拐角点开始突出

结论：低阶系数能够反映大体形状，高阶系数可以精确定义形状特征，少数傅里叶描述子携带了形状信息，能够反映边界的大略本质。

表示与描述



- 傅里叶描述符
  - ✓ 3）使用价值

（1）较少的傅里叶描述子（如4个），就可以获取边界本质的整体轮廓

（2）这些带有边界信息的描述子，可以用来区分明显不同的边界

表示与描述



- 傅里叶描述符
  - ✓ 4）优点

（1）使用复数作为描述符，对于旋转、平移、缩放等操作和起始点的选取不十分敏感。

（2）几何变换的描述子可通过对函数作简单变换来获得

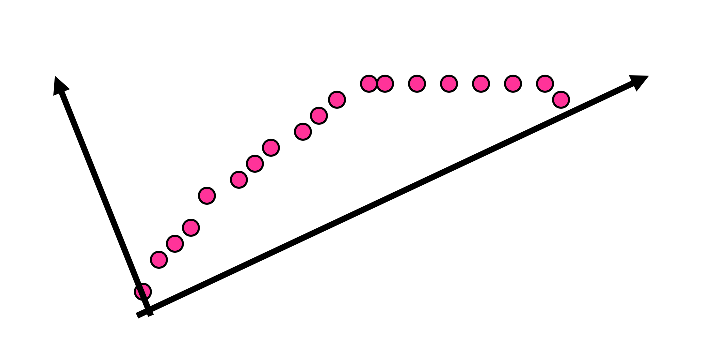
几何变换	傅里叶描述子
原形	$a(u)$
旋转	$a_r(u) = a(u)\ e^{j\theta}$
平移	$a_t(u) = a(u) + \Delta_{xy}\delta(u)$
缩放	$a_s(u) = \alpha a(u)$
起点	$a_p(u) = a(u)\ e^{-j2\pi k_0 u/N}$

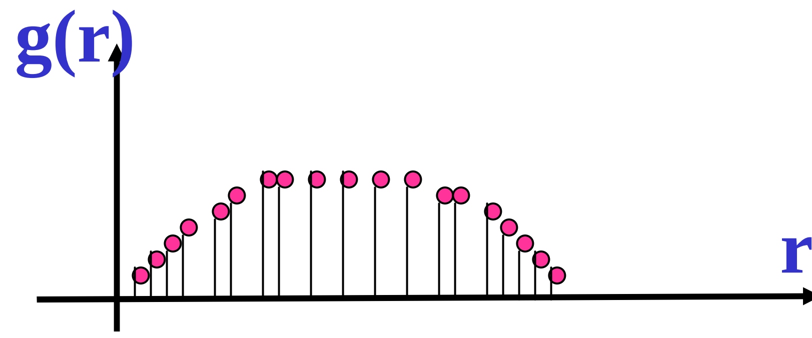
表示与描述



- 矩量
  - ✓基本思想：

    将描述形状的任务减少至描述一个一维函数，边界段和特征的形状可以用矩量来量化地描述
  - ✓矩量的定义：
    - 把边界当作直方图函数：g(r)





表示与描述



- 矩量
  - ✓ 矩量的定义：

$$\mu_n(r) = \sum_{i=1}^{L} (r_i - m)^n g(r_i)$$

其中 $$m = \sum_{i=1}^{L} r_i g(r_i)$$

这里L是边界上点的数目，$\mu_n(r)$是边界的矩量

表示与描述



- 矩量
  - ✓ 矩量的优点：
    - ➢ 实现是直接的
    - ➢ 附带了一种关于边界形状的“物理”解释
    - ➢ 对于旋转的不敏感性
    - ➢ 为了使大小比例不敏感，可以通过伸缩r的范围来将大小正则化。

表示与描述



- 概述
- 表示方法
- 边界描述子
- 关系描述子

表示与描述



- 关系描述子
  - ✓ 基本思想
  - ✓ 阶梯关系编码
  - ✓ 骨架关系编码
  - ✓ 方向关系编码
  - ✓ 内角关系编码
  - ✓ 树结构关系编码

表示与描述

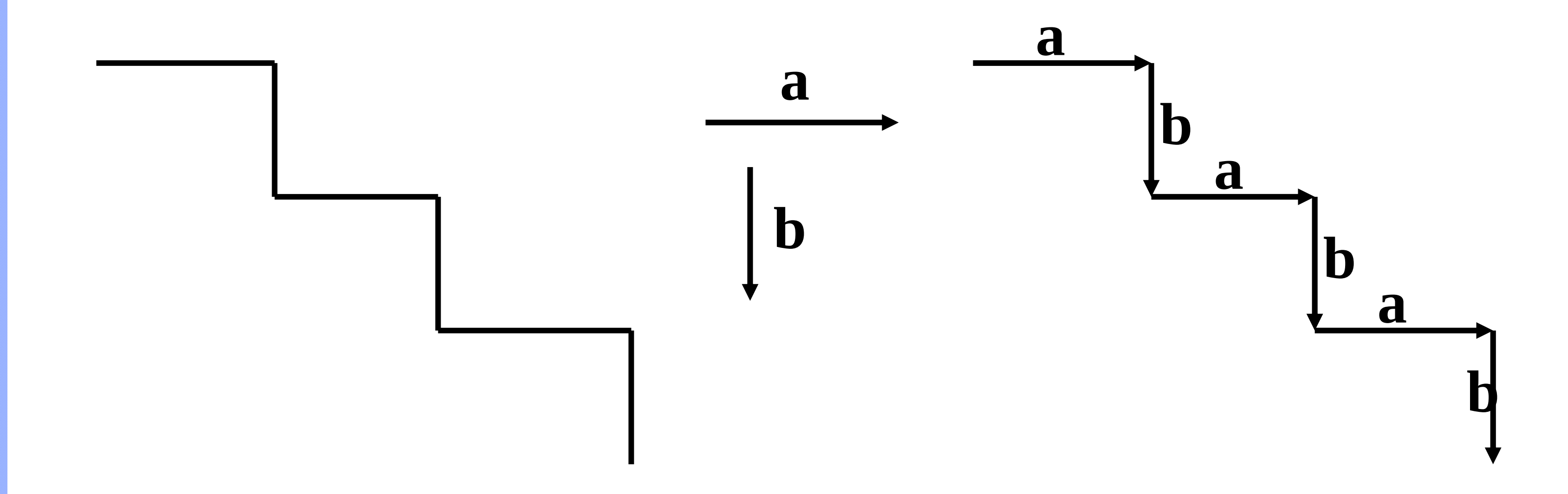


- 基本思想：
  - ✓ 通过挖掘各个成分之间的结构关系来描述边界
  - ✓ 图像中各个部分间的结构关系是二维的，而串是一维的，期望找到一种方法把二维关系转化为一维的串
  - ✓ 主导思想是考虑物体各个部分的连接线段

表示与描述



- 阶梯关系编码
  - ✓ 对于如下阶梯形边界，定义两个基本元素a, b



表示与描述



- 阶梯结构关系
  - ✓ 定义如下产生规则：

(1) S⟶aA 表明起始符S可以被图元a和变元A代替

(2) A⟶bS 表明变元A可以被b和S代替

(3) A⟶b 变元A可以只被b代替

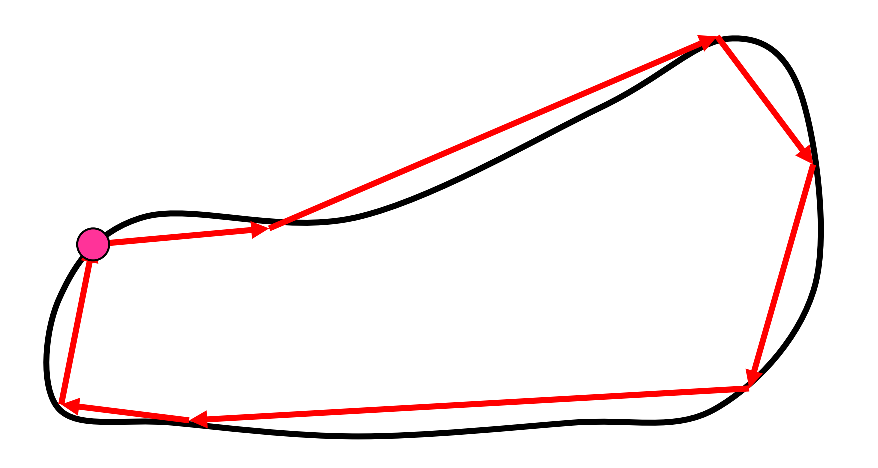
其中S、A是变量

举例：

表示与描述



- 骨架关系编码
  - ✓ 用有向线段来描述一个图像的各个部分（例如同构区域），这个线段是通过头尾连接等方法得到的。线段之间的不同运算代表了区域的不同组合。
  - ✓ 当图像的连通性可以通过首尾相接或其它连续的方式描述的时候，最适于使用这种串来描述。



表示与描述



- 骨架关系编码

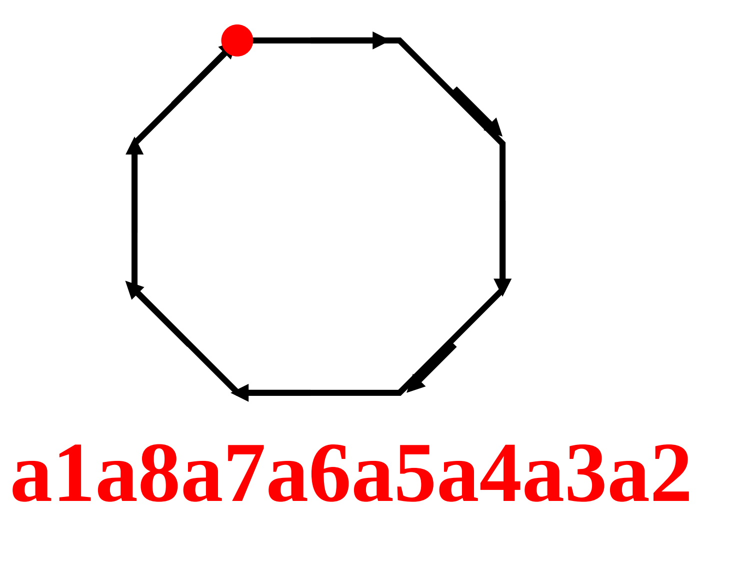
编码 c + a　c - a　a × b　a * b

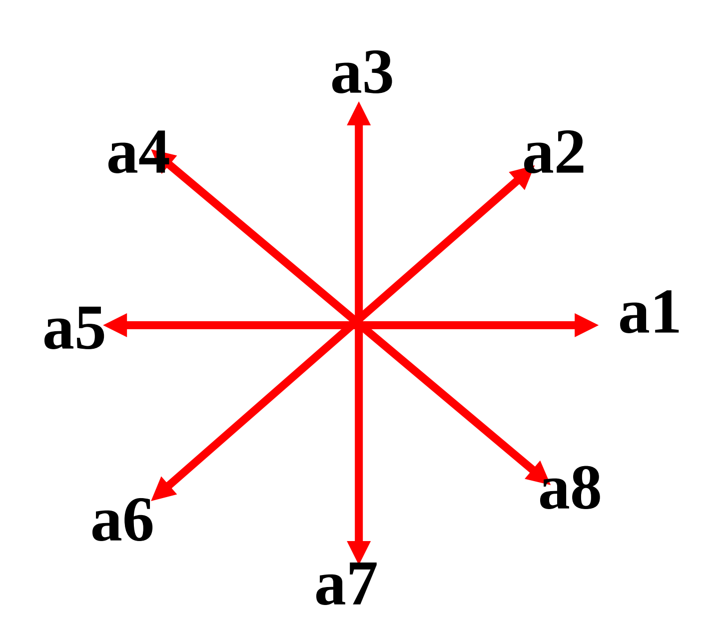
a + a + b + e + e + e + a

表示与描述



- 方向关系编码
  - ✓ 跟踪对象的边界，将跟踪得到的线段按照方向或长度来编码





表示与描述



- 内角关系编码
  - 根据内角角度范围不同，编码为8个符号

    即：　a1:0-45; a2:45-90;a3:90-135;...;
    a8:315-360

    举例：

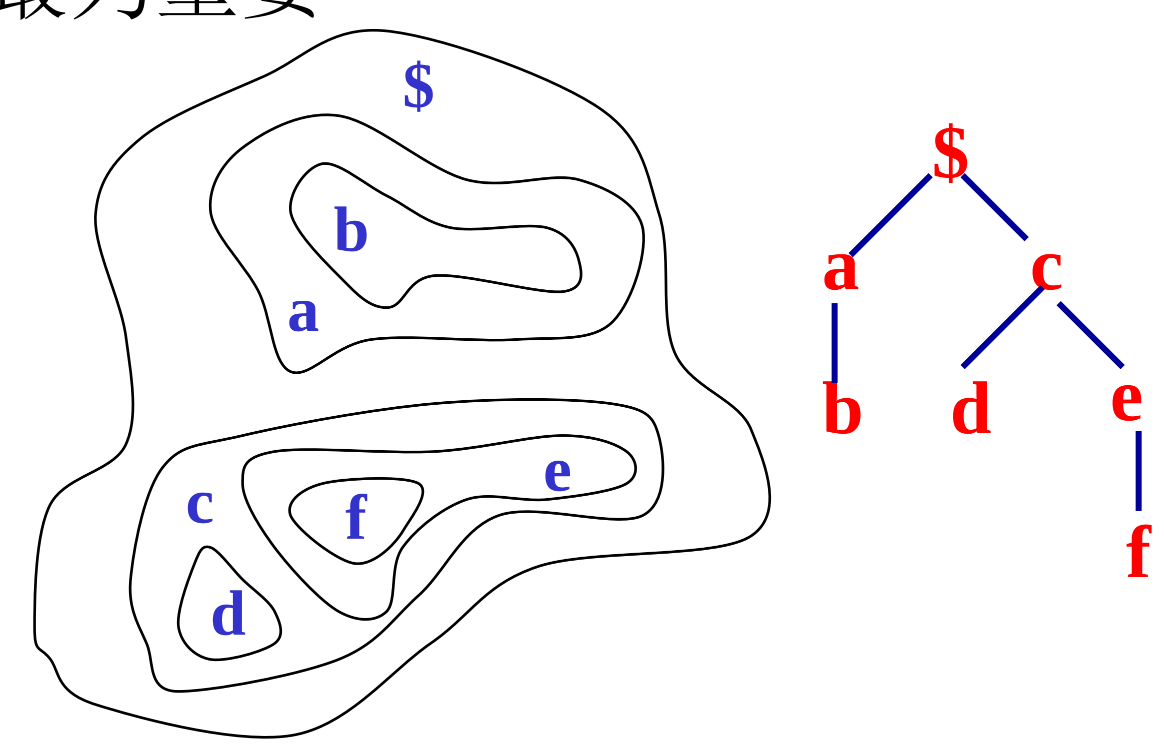
    a3a3a3a3a3a3a3a3　　　**a2a2a3a3**

表示与描述



- 树结构关系
  - ✓ 树结构中每个结点的意义和结点之间的关系最为重要

举例：





任何问题？